

MONITORIZAÇÃO E CARACTERIZAÇÃO DE ATIVIDADE FÍSICA

Sérgio Hélder da Silva Soares



Departamento de Engenharia Eletrotécnica

Mestrado em Engenharia Eletrotécnica e de Computadores

Área de Especialização em Telecomunicações

2014

Este relatório satisfaz, parcialmente, os requisitos que constam da Ficha de Disciplina de Tese/Dissertação, do 2º ano, do Mestrado em Engenharia Eletrotécnica e de Computadores

Candidato: Sérgio Hélder da Silva Soares, N° 1070312, 1070312@isep.ipp.pt

Orientação científica: Pedro Miguel Machado Soares Carvalho,

pedro.carvalho@inescporto.pt, pms@isep.ipp.pt Empresa: INESC Porto

Co-Orientação: Paula Maria Marques Moura Gomes Viana, pmv@isep.ipp.pt



Departamento de Engenharia Eletrotécnica

Mestrado em Engenharia Eletrotécnica e de Computadores

Área de Especialização em Telecomunicações

2014

Agradecimentos

Quero agradecer a orientação e atenção dispensada pela Engenheira Paula Viana, e pelo meu supervisor, Engenheiro Pedro Carvalho, pela forma tão especial de comunicar e transmitir os seus conhecimentos e ideias e por terem aberto os meus horizontes, mostrar-me que apesar de não se saber não quer dizer que não consiga.

Agradeço, profundamente aos meus pais, a minha base segura, que permitiram que tudo isto fosse possível. À minha irmã Sandra pela paciência, por me terem sempre apoiado e ajudado. À minha sobrinha Lara por toda a alegria e brincadeira que ajudaram nos momentos menos bons. À minha avó, por todo o carinho e ensinamentos que me deu ao longo da vida. À minha namorada Ana, por ter muita paciência e insistência para acabar o Mestrado, por toda a felicidade, apoio e motivação para concretizar este projeto.

Aos meus amigos, pela força, amizade e por todos os momentos de convívio que partilhámos ao longo destes anos.

Esta dissertação foi desenvolvida também no contexto e em colaboração com os projetos: Media Arts and Technologies (MAT), NORTE-07-0124-FEDER-000061, financiado pelo Programa Operacional Regional do Norte (ON.2 – O Novo Norte), sob o Quadro de Referência Estratégica Nacional (QREN), através do Fundo Europeu de Desenvolvimento Regional (FEDER), e por fundos nacionais através da agência de financiamento Portuguesa, Fundação para a Ciência e a Tecnologia (FCT); QREN 33910 ARENA, um projeto de I&D financiado pelo Fundo Europeu de Desenvolvimento Regional (FEDER) através do Programa Operacional Regional do Norte (ON.2 – O Novo Norte) como parte do Quadro de Referência Estratégica Nacional (QREN), e gerido pelo IAPMEI - Agência para a Competitividade e Inovação, I.P”.

Resumo

A monitorização da atividade física é um tema que tem adquirido cada vez mais importância. Tal deve-se ao crescente sedentarismo da população em geral e adquirindo níveis muito elevados de importância devido a vários fatores como por exemplo o enorme crescimento tecnológico e menor tempo de lazer. Cada vez mais a população tem a tendência de substituir atividades como uma simples caminhada para o trabalho ou escola por algum tipo de tecnologia que reduz o consumo energético do corpo, sendo paradigmático o uso (excessivo) de viaturas automóveis. Em consequência da escassez de atividade física, doenças como a obesidade e problemas cardíacos têm vindo a aumentar nas várias faixas etárias, mas assume uma particular relevância em crianças.

Nas últimas décadas têm aumentado as iniciativas de investigação com o objetivo de compreender os fatores que afetam a prática de atividade física para posteriormente a potenciar. Existem diversos métodos contudo, destaca-se preferencialmente os de observação direta, com observadores presentes. No entanto estes apresentam algumas limitações. Consequentemente são necessários esforços de investigação adicionais e novas técnicas ou metodologias.

Nesta dissertação pretende-se contribuir ativamente para a investigação na área da promoção de atividade física através da utilização de vídeo, com uma análise realizada sobre dois pontos principais. Primeiro são analisados métodos do estado de arte que requerem a presença de observadores e de que forma a captura de vídeos pode ser utilizada como alternativa ou complemento. De seguida, é realizado um estudo e avançada uma proposta inicial para utilizar mecanismos de processamento e classificação automática da atividade em alternativa ao observador humano.

Palavras-Chave

Caracterização e monitorização da atividade física; Segmentação; Anotação;
Classificação automática

.

Abstract

Monitoring of physical activity is a topic that has gained increasing importance. This is due to the increased physical inactivity in the general population and has acquired very high levels due to various factors such as the huge technological growth and less leisure time. Increasingly, people have a tendency to substitute activities as a simple walk to work or school for some kind of technology that reduces energy consumption of the body, being paradigmatic use (excessive) of the car. In consequence of the scarcity of physical activity, obesity and diseases such as heart problems are increasing in all age groups, but assumes particular relevance in children.

In the last decades have increased research initiatives with the aim of understanding the factors that affect physical activity to potentiate later. Several methods have been proposed, highlighting the direct observation, with observers being preferred. However these present some limitations. Consequently required additional efforts and new research techniques or methodologies.

In this dissertation is intended to actively contribute to research in the area of physical activity promotion through the use of video, with an analysis of two major objectives. First analysed the state of the art methods that require the presence of observers and in what way capture videos can be used as alternative or supplement. Next a study is made and advanced a proposal to use processing mechanisms and automatic classification of activity as an alternative to the human observer

Keywords

Characterization and monitoring of physical activity; segmentation; annotation; Automatic classification.

Índice

AGRADECIMENTOS	I
RESUMO	III
ABSTRACT	V
ÍNDICE	VII
ÍNDICE DE FIGURAS	IX
ÍNDICE DE TABELAS	XI
ACRÓNIMOS.....	XIII
1. INTRODUÇÃO	1
1.1. CONTEXTUALIZAÇÃO	1
1.2. OBJETIVOS	2
1.3. CONTRIBUIÇÕES	3
1.4. ESTRUTURA DA DISSERTAÇÃO	3
2. MÉTODOS E TECNOLOGIAS PARA CARACTERIZAÇÃO DA ATIVIDADE FÍSICA	5
2.1. OBSERVAÇÃO DIRETA.....	5
2.1.1. <i>The Fargo Activity Time sampling (FATS)</i>	7
2.1.2. <i>Activity Patterns and Energy Expenditure (APEE)</i>	8
2.1.3. <i>Children's Physical Activity Form (CPAF)</i>	8
2.1.4. <i>Children's Activity Rating Scale (CARS)</i>	8
2.1.5. <i>System for Observation Fitness Instructor Time (SOFIT)</i>	8
2.1.6. <i>System to Observe Physical and Leisure Activity in Youth (SOPLAY)</i>	8
2.1.7. <i>System for Observing Play and Recreation in Communities (SOPARC)</i>	9
2.1.8. <i>SOFIT Vs. SOPLAY</i>	9
2.2. TECNOLOGIA DE APOIO À CARACTERIZAÇÃO DA ATIVIDADE FÍSICA.....	10
2.2.1. <i>Sistema de posicionamento global</i>	10
2.2.2. <i>Acelerómetros</i>	11
2.2.3. <i>Vídeo</i>	12
3. PROCESSAMENTO AUTOMÁTICO DE INFORMAÇÃO	15
3.1. SEGUIMENTO AUTOMÁTICO DE PESSOAS	15
3.2. MACHINE LEARNING E DATA MINING	21
3.3. BIBLIOTECAS COM SUPORTE PARA <i>MACHINE LEARNING</i>	28
3.3.1. <i>Waikato Enviroment for Knowledge Analysis - Weka</i>	28
3.3.2. <i>Matrix Laboratory – MATLAB</i>	28
3.3.3. <i>Bibliotecas Machine Learning no Python</i>	29

3.3.4.	<i>OpenCV</i>	29
3.3.5.	<i>Linguagem R</i>	30
3.3.6.	<i>LIBSVM</i>	30
4.	PREPARAÇÃO E CARACTERIZAÇÃO DOS DADOS	33
4.1.	CONJUNTO DE DADOS	33
4.2.	DESCRIÇÃO METODOLOGIAS DE CARACTERIZAÇÃO E ANOTAÇÃO.....	35
4.3.	CARACTERIZAÇÃO DOS DADOS DA OBSERVAÇÃO DIRETA.....	37
4.3.1.	<i>Comparação das duas Experiências de anotação</i>	39
4.3.2.	<i>Classificação Movimento a partir de um painel de Observadores</i>	44
5.	ABORDAGEM E TESTE PARA CARACTERIZAÇÃO AUTOMÁTICA.....	47
5.1.	SIMULAÇÃO DE UM ALGORITMO DE SEGUIMENTO.....	47
5.2.	DESENVOLVIMENTO DA APLICAÇÃO	48
5.3.	DEFINIÇÃO DO CLASSIFICADOR AUTOMÁTICO	50
5.4.	SELEÇÃO DO CONJUNTO DE DADOS DE TREINO E TESTE	51
5.5.	TREINO E TESTE DO SISTEMA DE CLASSIFICAÇÃO.....	52
5.6.	ANÁLISE DOS RESULTADOS	55
5.7.	COMPARAÇÃO ENTRE A CLASSIFICAÇÃO AUTOMÁTICA E A CLASSIFICAÇÃO TRADICIONAL	57
6.	CONCLUSÕES E DESENVOLVIMENTOS FUTUROS.....	63
	REFERÊNCIAS DOCUMENTAIS	65

Índice de Figuras

Figura 1 – Atividade Física	1
Figura 2 - Diferentes classes de movimento.	7
Figura 3 - Imagem capturada (a) e sua segmentação (b) do projeto CAVIAR [25].....	16
Figura 4 – Diferentes tipos de ocultação [25].	17
Figura 5 – Diferentes representações de um objeto para o seguimento [25].....	18
Figura 6 – Ilustração do método <i>predict-match-update</i> [25].	20
Figura 7 – Diagrama típico de um problema de aprendizagem (adaptado de [35]).	22
Figura 8 – Modelo de classificação.....	23
Figura 9 – Exemplo Árvore de decisão [36].	24
Figura 10 – Possível mapeamento de diferentes classes pelo SVM em uma classificação Linear [44].	25
Figura 11 – Exemplo duas classes onde é impossível o SVM separá-las [40].....	26
Figura 12 – Exemplo da transformação realizada por um <i>kernel</i> [42].	26
Figura 13 – Exemplo de criação modelo de treino.....	31
Figura 14 – Exemplo de classificação e percentagem de precisão do modelo.....	31
Figura 15 - Perspetiva da sequência de um grupo.....	34
Figura 16 – Perspetiva da sequência de um grupo.	34
Figura 17 – Excerto formulário de anotação	36
Figura 18 – Anotações sem paragens no conjunto dos vídeos	38
Figura 19 - Anotações com paragens no conjunto dos vídeos	39
Figura 20 – Gráfico de confusão por tipo de movimento para o Observador 1 sem paragens....	41
Figura 21 – Gráfico de confusão por tipo de movimento para o Observador 1 com paragens ...	41
Figura 22 – Gráfico de confusão por tipo de movimento para o Observador 2 sem paragens....	42
Figura 23 - Gráfico de confusão por tipo de movimento para o Observador 2 com paragens....	42
Figura 24 - Gráfico de confusão por tipo de movimento para o Observador 3 sem paragens	43
Figura 25 - Gráfico de confusão por tipo de movimento para o Observador 3 com paragens....	43
Figura 26 – Processo de criação da classificação automática	47
Figura 27 – Exemplo da aplicação.	49
Figura 28 – Excerto do ficheiro output XML.....	49
Figura 29 – Gráfico cruzamento melhores classificadores.....	57
Figura 30 – Gráfico de tendência acumulada da média dos dois classificadores nos dois modelos de referência.	60

Índice de Tabelas

Tabela 1 – Exemplo de uma tabela de confusão	27
Tabela 2 – Propriedades das sequências de vídeo.....	35
Tabela 3 – Caracterização dados Anotação sem paragens	37
Tabela 4 - Caracterização dados Anotação com paragens	38
Tabela 5 – Dados comparativos das duas fases de anotação.....	39
Tabela 6 – Tempo realizado para cada tipo de anotação.....	40
Tabela 7 – Confiança por movimento e tipo de observação.	45
Tabela 8 – Seleção dados de teste e treino	51
Tabela 9 – Resultados da classificação com <i>kernel</i> Linear	53
Tabela 10 - Resultados da classificação com <i>kernel não</i> Linear.....	54
Tabela 11 - Resultados da classificação com <i>kernel não</i> Linear com normalização	55
Tabela 12 – Cruzamento dos melhores resultados de todos classificadores	56
Tabela 13 – Comparação do acerto entre Observação Tradicional e Automática com os dois modelos de referência.....	59
Tabela 14 – Média do acerto dos classificadores com os dois modelos de referência.....	60

Acrónimos

APEE	–	Activity Patterns and Energy Expenditure
CARS	–	Children’s Activity Rating Scale
CPAF	–	Children’s Physical Activity Form
FATS	–	Fargo Activity Time Sampling
GIS	–	Geographic Information Systems
GPS	–	Global Positioning System
GNSS	–	Global Navigation Satellite System
MATLAB	–	Matrix Laboratory
SOFIT	–	System for Observation Fitness Instructor Time
SOPARC	–	System for Observing Play and Recreation in Communities
SOPLAY	–	System to Observe Physical and Leisure Activity in Youth
WEKA	–	Waikato Environment for Knowledge Analysis
XML	–	eXtensible Markup Language

1. INTRODUÇÃO

1.1. CONTEXTUALIZAÇÃO

Os avanços científicos na área da medicina têm permitido melhorar a resposta a um conjunto de doenças contribuindo para o aumento progressivo da esperança e qualidade de vida. Para além destes avanços científicos, está também comprovado que a atividade física regular tem um impacto importante na prevenção de um número significativo de doenças crónicas, da obesidade cuja incidência tem vindo a aumentar devido ao estilo de vida cada vez mais sedentário da população em geral [1]. Devido a estes fatores, tem aumentado o interesse no estudo da monitorização e caracterização da atividade física de maneira a auxiliar a combater estas tendências da população. Os estudos revelam que a atividade física regular na população é influenciada por motivos culturais, sociais e ambientais em que estas estão integradas [2][6]. Este impacto significativo da atividade física na saúde tem contribuído para que outras ciências, para além da ciência médica, desenvolvam soluções que contribuam para avaliar o impacto destas ações e melhorar o seu desempenho tendo como objetivo final a promoção de atividade física adequada a cada indivíduo.



Figura 1 – Atividade Física

A caracterização da atividade física é importante, pois permite estudar a atividade física em diferentes faixas etárias. A caracterização não se limita à análise de desportos abrangendo também o dia-a-dia das pessoas, como objetivo de perceber quais as melhores metodologias através da sobreposição de dados que são adquiridos ao longo da atividade, ou rotina diária. A caracterização da atividade física pretende, através da observação dos

movimentos do ser humano, retirar dados que permitam perceber quais os problemas ou os movimentos mais comuns na pessoa em estudo, para assim poder ajudar a melhorar a sua qualidade de vida através de exercício físico adequado. A caracterização pode ser realizada através de observação direta, com um observador presente ou então com o apoio de tecnologias como o GPS ou acelerómetros.

A observação direta da atividade física é um processo complexo que envolve normalmente a observação direta e a realização de questionários e registos diários. Consequentemente, a sua execução é um processo demorado e com custos com pessoal, relevantes. Além disso, os métodos atuais apresentam limitações pelo fato de todo o processo de caracterização ser realizado em tempo real, com a observação e registo realizadas no momento. Isto pode resultar em erros na anotação, quer por limitação do observador, quer por limitações do ambiente/local em que está a ser realizada a atividade física. Com o intuito de melhorar este processo de recolha e análise de informação, e de modo a reduzir os erros associados aos fatores atrás descritos, a utilização de tecnologias de informação e comunicação podem contribuir de forma significativa para tornar este processo mais eficiente e objetivo. Entre estes, destacam-se a captura e análise de sinais de vídeo, a utilização de dispositivos móveis de medição e os desenvolvimentos de aplicações que facilitem a análise dos dados recolhidos.

Com a captura de imagens de vídeo é possível o desenvolvimento de métodos automáticos para a deteção, seguimento e caracterização das pessoas que se encontram no ambiente local, tornando a monitorização da atividade física mais flexível e, idealmente, objetiva. No entanto, a caracterização através da análise automática de imagens vídeo apresenta desafios importantes relacionados, como por exemplo a qualidade de imagem, a luminosidade e o local de instalação das câmaras de captação, a complexidade exigida a algoritmos de processamento automático. Qualquer um destes fatores pode causar dificuldades no processo automático de classificação/caracterização da atividade física.

1.2. OBJETIVOS

Pretende-se com esta dissertação apresentar uma proposta para o estudo da caracterização da atividade física recorrendo à utilização de vídeo. Assim, os objetivos desta dissertação são os seguintes:

1. Realizar uma análise e caracterização dos métodos do estado da arte;

2. Analisar a utilização de vídeo em conjunto com a aplicação de técnicas do estado da arte;
3. Definir e preparar uma abordagem para a classificação automática da atividade física.

1.3. CONTRIBUIÇÕES

As principais contribuições desta dissertação podem ser resumidas da seguinte forma:

- Preparação de um conjunto de dados de teste e avaliação, incluindo a geração de informação de referência de acordo com o output: da observação direta; de algoritmos de seguimento automático utilizando vídeo.
- Análise e caracterização de métodos de observação direta com e sem recurso a vídeo;
- Proposta de uma abordagem para classificação automática de atividade física utilizando vídeo e análise do impacto da utilização destas técnicas;
- Colaboração na elaboração de um artigo científico: Pedro Silva, Sérgio Soares, Paula Viana, Jorge Mota, Pedro Carvalho, “Reviving direct observation methods for physical activity behavior”, *Journal of Physical Activity & Health*, (submetido).

1.4. ESTRUTURA DA DISSERTAÇÃO

O trabalho encontra-se dividido em 5 capítulos. No primeiro capítulo é apresentado uma breve introdução ao tema, a problemática que levou à sua realização, o contexto desta dissertação e os objetivos definidos. No capítulo 2, é concretizado um resumo do estado da arte que inclui uma análise dos métodos e técnicas existentes para a caracterização da atividade física. Inclui também um estudo de tecnologias adjacentes que possam contribuir ou ser usados para a caracterização da atividade física. No capítulo 3, são descritas algumas técnicas utilizadas no processamento automático para deteção e seguimento de pessoas, técnicas de seguimento. São também descritas sumariamente técnicas de *machine learning e data Mining*, descrição de modelos de classificação e algumas aplicações que usam bibliotecas de *machine learning*. O capítulo 4, está dividido em diferentes secções. A primeira está relacionada com a seleção de sequências típicas em ambientes reais. Na segunda, com ajuda de observadores treinados na caracterização,

é descrita uma caracterização por observação direta das sequências selecionadas, para assim poder ter um ponto de comparação com a caracterização automática e realizar uma análise sobre as limitações encontradas por este método e uma última etapa é ilustrado o programa de realização de anotações em vídeos.

No capítulo 5, é realizada uma posposta de um classificador automático e uma descrição de todos os passos para a construção de um classificador automático, características e classes usadas para a construção de um modelo de classificador. São apresentados todos os testes e seus resultados, comparam-se os resultados obtidos com os da observação tradicional. No último capítulo, 6, são resumidas as principais conclusões e perspectivas futuros desenvolvimentos.

2. MÉTODOS E TECNOLOGIAS PARA CARACTERIZAÇÃO DA ATIVIDADE FÍSICA

Neste capítulo é apresentado um estudo sobre dois pontos principais de forma ajudar a entender o funcionamento da caracterização da atividade física: métodos e formas de caracterização e descrição da atividade física; as tecnologias e abordagens na área computacional, nomeadamente, detecção e seguimento de objetos em vídeo.

O avanço tecnológico tem sido umas das principais causas do sedentarismo. No entanto, estas mesmas tecnologias podem ser usadas para combater esse fenómeno e contribuir de forma significativa para melhorar o impacto que o exercício físico pode ter sobre um indivíduo. Atualmente, existem várias tecnologias usadas para a caracterização da atividade física tais como: aplicação de acelerómetros, do sistema de posicionamento global (GPS), do sistema de informação geográfica (GIS), *Smartphones* e sistemas de processamento automático baseados em imagens ou vídeo.

A informação que é possível recolher através de tecnologia permite caracterizar a atividade física regular do ser humano e também, pode ser utilizada para determinar a eficácia de um conjunto de ações na saúde do mesmo. Apesar da proliferação da tecnologia, o método mais utilizado é ainda a observação direta, que consiste em observar diretamente no local em tempo real sem tecnologias inerentes.

2.1. OBSERVAÇÃO DIRETA

Este método possibilita um observador treinado avaliar diretamente, o comportamento dos intervenientes e a sua atividade física, assim como o contexto em que ocorreu. Observação direta consiste em uma pessoa ou um grupo serem observados durante um

determinado período de tempo por um observador e fazer uma anotação da atividade física do observado durante o mesmo tempo de observação.

Uma das limitações associadas à observação direta está associada à análise e classificação dos dados em tempo real, e à necessidade de treinar os observadores para uma avaliação imparcial. Ao ser realizada por diferentes observadores, cria questões relacionadas com a subjetividade da caracterização pelo facto de não existir uma metodologia realmente única e objetiva. É um método que se torna subjetivo isto porque ao ser realizado por pessoas num espaço temporal curto e irrepetível, onde existem parâmetros que se tornam difíceis de medir objetivamente através da visão.

Para uma análise adequada a atividade física deve ser natural. No entanto, a presença do observador tende a influenciar e por vezes condicionar o comportamento. Este é um ponto fulcral na caracterização da atividade física: a possibilidade de caracterizar sem influenciar a sua realização. Métodos que se baseiam no uso de dispositivos invasivos (no sentido que tem de ser usados pelos observados), como o caso de acelerómetros e dispositivos GPS, podem ter um impacto negativo neste requisito. A observação direta com recurso a sistemas de processamento de vídeo pode neste caso apresentar algumas vantagens, pois retiram qualquer tipo de limitação na relação observador/sujeito e é não invasivo. Com o avanço de tecnologias de seguimento com base em vídeos, estas poderão ser de grande utilidade e aplicadas para o melhoramento da monitorização e caracterização da atividade física. A introdução destas novas tecnologias contribuem para uma maior resolução temporal das observações e diminui a carga no observador humano. Além disso, com o aumento da automação melhorou a objetividade do processo e aumentou o tipo e a diversidade dos dados recolhidos.

A caracterização da atividade física exige uma análise dos tipos de movimentos podendo incluir entre outras, a velocidade de movimentação do atleta ou a sua posição em relação a um referencial, sendo assim possível identificar os movimentos da pessoa, como por exemplo saltar no lugar ou então um salto durante uma corrida.

Na literatura, tipicamente a classificação do comportamento de uma pessoa inclui três tipos de classes de atividade física. Essas três classes são de uma maneira geral: sedentário, moderado e vigoroso, do qual dependendo do tipo de caracterização podem existir subclasses como: deitado, sentado, em pé, caminhar.

O comportamento sedentário está associado a situações em que a pessoa tem “pouco movimento” ou está parada; Compreende as subclasses estar em “pé”, “sentado” e “deitado”. A classe de comportamento moderado está associado à existência de algum movimento. É uma classe que separa a pessoa estar parada e quando entra num movimento vigoroso, incluindo-se neste estado o comportamento onde existe algum tipo de movimento, (como andar), mas termina quando esse movimento se torna em algo mais intenso, como correr. Podem incluir-se diversas subclasses “marcha lenta” e “marcha”. Por fim, comportamento vigoroso está associado a um grau de movimento elevado onde a pessoa se movimenta constantemente, compreendendo as subclasses “correr” e “saltar” (Figura 2).

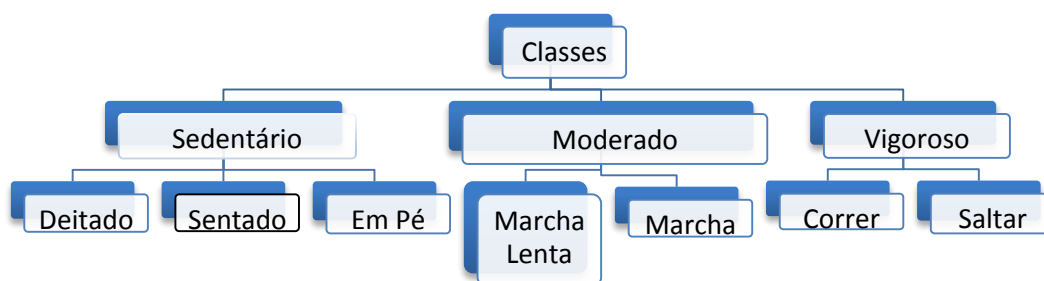


Figura 2 - Diferentes classes de movimento.

Existem várias técnicas que tem vindo a ser desenvolvidas ao das últimas décadas para a caracterização da atividade física por observação direta tais como [7]:

- The Fargo Activity Timesampling (FATS);
- Activity Patterns and Energy Expenditure (APEE);
- Children’s Physical Activity Form (CPAF);
- Children’s Activity Rating Scale (CARS);
- System for Observation Fitness Instructor Time (SOFIT);
- System to Observe Physical and Leisure Activity in Youth (SOPLAY);
- System for Observing Play and Recreation in Communities (SOPARC).

2.1.1. THE FARGO ACTIVITY TIME SAMPLING (FATS)

Desenvolvido por Klesges, 1984, foi a primeira tentativa de desenvolver um instrumento para descrever a atividade física das crianças e o comportamento dos pais. O procedimento em FATS consiste na observação durante um intervalo de tempo seguindo

de registo, ou seja, é observado o comportamento do grupo ou pessoa por 10 segundos e em seguida é registado o comportamento durante 10 segundos [7].

2.1.2. **ACTIVITY PATTERNS AND ENERGY EXPENDITURE (APEE)**

Neste método desenvolvido por Epstein em 1984, observado a atividade física por um intervalo de tempo de 15 segundos e classifica o nível de atividade numa das 5 categorias [7].

2.1.3. **CHILDREN'S PHYSICAL ACTIVITY FORM (CPAF)**

O CPAF foi concebido para medir a atividade física durante as aulas de educação física, utilizando intervalos de um minuto de observação e anotação sendo desenvolvido por O'Hara, 1989 [7].

2.1.4. **CHILDREN'S ACTIVITY RATING SCALE (CARS)**

Desenvolvido por Puhl em 1990, o CARS consiste numa escala de cinco níveis: em descanso, atividade nível baixo, atividade de nível médio, atividade de nível médio-alto e atividade de nível vigoroso, foi concebido para categorizar a intensidade das atividades físicas e criar uma diferenciação entre diversos níveis de gasto energético em crianças pequenas. O sistema utiliza observações com intervalos de 1 minuto para anotação [7].

2.1.5. **SYSTEM FOR OBSERVATION FITNESS INSTRUCTOR TIME (SOFIT)**

Foi desenvolvido por McKenzie em 1991, SOFIT consiste na observação da atividade física em intervalos de 10 segundos e 10 segundos para registar pessoas individualmente. Neste método são usadas cinco classes: deitado, sentado, em pé, caminhar e vigoroso [7].

2.1.6. **SYSTEM TO OBSERVE PHYSICAL AND LEISURE ACTIVITY IN YOUTH (SOPLAY)**

Este método desenvolvido por McKenzie em 2000 foi criado com o objetivo de permitir quantificar a atividade física em grupo, assim como o contexto ambiental em que a atividade se desenvolve. É baseado em observações onde é realizado uma observação em varrimento, (observador dentro do seu campo de visão realizar a observação da esquerda para a direita sem voltar atrás no campo de visão), da atividade física de um determinado

grupo. O SOPLAY permite indicar o número de atletas, sexo e define três categorias: sedentário, moderado/andar, vigoroso [7].

2.1.7. ***SYSTEM FOR OBSERVING PLAY AND RECREATION IN COMMUNITIES (SOPARC)***

Desenvolvido a partir de SOPLAY por Brown em 2006 e tem como objetivo capturar a atividade física e informação contextual em pessoas que frequentam ambientes comunitários, como parques e áreas de recreação. Uma área é varrida de forma sistemática, da esquerda para a direita e os níveis de atividade e tipos são registradas para diferentes sexos, em simultâneo com a informação contextual [7].

No âmbito desta tese são tomadas como referência as técnicas SOFIT e SOPLAY isto porque são atualmente os métodos mais usados pelos investigadores especializadas na área. De seguida, é apresentada uma caracterização mais específica de cada uma.

2.1.8. **SOFIT Vs. SOPLAY**

O SOFIT fornece informação detalhada de um grupo reduzido de pessoas em que cada pessoa é observada durante um determinado período de tempo. Apesar de vários estudos indicarem que pode ser usada de forma fiável em diversos contextos. O SOFIT apresenta limitações sobretudo no que diz respeito ao facto das observações serem realizadas pessoa a pessoa, o que significa que quanto maior for o número de avaliados, maior será o tempo que vai exigir, tal facto, o SOPLAY foi criado para combater essa limitação.

O SOPLAY é baseado em amostragens temporais onde é realizado uma observação da atividade física de um determinado grupo, ou seja, a técnica consiste em fazer uma observação de toda a área de análise, da esquerda para a direita. O output consiste na indicação do número de atletas e é quantificado separadamente em três categorias. No tempo em que cada observação é realizada, tem de haver registo por parte do observador, sendo o intervalo de cada observação flexível, ou seja, o observador faz a observação e após essa observação tem um tempo de registo antes de realizar outra observação. O tempo mais utilizado para cada observação situa-se normalmente entre 10 a 30 segundos. Quando existe uma necessidade de um estudo mais detalhado o tempo de observação pode ir até 2 minutos.

O SOPLAY é aplicado para identificar pessoas presentes ou envolvidos em diferentes atividades físicas, assim como os fatores contextuais associados, como por exemplo a existência de condições para a observação ser realizada, condições meteorológicas, equipamentos e local. A atividade física e os fatores contextuais da área são codificados, existindo diferentes classificações para cada ação e comportamento. São também realizadas análises separadas para rapazes, raparigas e para diferentes equipas, (se for o caso), de modo a permitir a comparação de diferentes grupos. O SOPLAY também indica o tipo de atividade física que está a ser realizada. Sendo uma técnica de observação direta, o SOPLAY tem limitações pois exige que o observador tenha um elevado nível de treino. Cada observador tem diferentes tempos para fazer a observação do grupo; e esporadicamente dois observadores terão o mesmo tempo de observação, o que cria problemas de inconsistência nos dados. Uma outra limitação consiste na utilização desta técnica em situações onde existe um número muito grande de sujeitos, sendo quase impossível não haver repetição da contagem dos sujeitos envolvidos na observação, pois como é realizado um varrimento do campo de visão do observador, na existência de uma quantidade grande de observados pode levar a uma contagem errada do número de pessoas.

2.2. TECNOLOGIA DE APOIO À CARACTERIZAÇÃO DA ATIVIDADE FÍSICA

2.2.1. SISTEMA DE POSICIONAMENTO GLOBAL

Para uma melhor compreensão da natureza da atividade física (ou inatividade) são exploradas todas as diferentes tecnologias que possam ajudar nessa caracterização. O GPS, (*Global Position System*), tem sido uma das tecnologias aplicadas pelos investigadores, pois tem a potencialidade de fornecer informações de localização. O Sistema de Posicionamento Global é atualmente o único *Global Navigation Satellite System*, (GNSS), totalmente funcional. Existem atualmente 24 satélites GPS na órbita da Terra a transmitir sinais para recetores de GPS, os quais determinam a localização, direção e velocidade do recetor, sendo por isso uma ferramenta importante para a navegação. O GPS também fornece uma referência temporal precisa [21].

A posição de um recetor GPS é calculada através da medição da distância entre o recetor e três ou mais satélites GPS. Cada satélite está equipado com um relógio atómico. Quando conectado pela primeira vez, os dispositivos GPS têm um período de

inicialização, durante a qual adquirem sinais dos satélites e sincronizam o relógio do GPS com o relógio atômico do satélite. Os dispositivos GPS usam a triangulação para determinar a posição do recetor, velocidade e altitude. Apesar da constante evolução desta tecnologia, o GPS apresenta algumas limitações no cálculo de posições dentro de edifícios, o que limita a sua utilização para a caracterização da atividade física em ambientes indoor [21].

Algumas das vantagens que se podem destacar na utilização de GPS [21]:

- São dispositivos portáteis (tamanho leve e pequeno);
- Efetuam medição contínua e permitem fácil acesso dos dados obtidos;
- O livre acesso dos satélites GPS em qualquer parte do mundo;
- Custo razoável;
- Os dados podem ser armazenados e subseqüentemente recuperados, se necessário;
- Técnica pode ser usada de forma independente para validar as medições de velocidade do andar e correr por outras técnicas.

Apesar das potenciais vantagens descritas, a cobertura dos satélites em algumas partes do mundo é deficiente resultado em uma disponibilidade limitada de acesso. A utilização do GPS para fazer a monitorização da posição de pessoas realizando atividades físicas, como jogos de futebol ou em desportos em que existe grande variação da atividade física, demonstrou que o GPS pode ser utilizado para monitorizar com precisão a posição e a velocidade do atleta em tempo real, sendo por isso uma ferramenta útil na avaliação do desempenho desportivo. No entanto não existem muitos estudos de aplicação do GPS para caracterizar a atividade física sem ser em desportos [21][22].

Uma outra limitação associada à utilização destes dispositivos para caracterização da atividade física, está relacionados com o facto de ser um processo invasivo, isto porque, a pessoa a ser observada sabe que está a ser estudado devido ao usar um dispositivo e por isso, pode alterar o seu comportamento natural na atividade física, pois este sabe que está a ser avaliada e como consequência os dados retirados não serão precisos e naturais.

2.2.2. ACELERÓMETROS

O acelerómetro é outro dispositivo que pode ser utilizado para a caracterização da atividade física. Pode ser aplicado em diferentes partes do corpo do sujeito em estudo e permite fazer a medição do movimento do corpo através da aceleração ao longo de um

ou mais eixos. As suas medições são realizadas nos três planos ortogonais e os dados podem ser guardados para análise posterior. Os acelerómetros usam vários tipos de tecnologias como cristais piezoelétrico, fontes micromecânicas, etc.[23]. Os dados podem conter múltiplas medições nos diferentes planos numa única área, permitindo assim existir diferentes planos de movimento a ser capturados [23][24].

Os acelerómetros têm sido utilizados em diferentes casos e provou-se que os seus dados eram confiáveis em diferentes tipos de atividade física e em diferentes sujeitos (crianças e adultos). No entanto, é importante referir que o número de medições depende do tipo de sujeito e a atividade que vai ser monitorizada. Durante o desenvolvimento do acelerómetro provou-se que a sua utilização tem maior consistência na caracterização da atividade física para atividades em que existam grandes movimentações, (andar, correr). No entanto têm sido desenvolvidos novas equações para melhorar a consistência em atividades como andar de bicicleta [23][24].

Em comparação com outros dispositivos, o acelerómetro tem a vantagem de distinguir entre o andar e a corrida em diferentes tipos de terreno. No entanto, tal como todas as tecnologias de caracterização da atividade física tem as suas limitações. Como o GPS é um método invasivo pois a pessoa tem de usar sempre um dispositivo. Além disso para que os dados adquiridos se tornem precisos e confiáveis, tem que existir um grande número de dados, de vários dias de atividade física [23][24].

2.2.3. VÍDEO

A utilização de vídeos na caracterização da atividade física é um método que em comparação com outros é menos invasivo, já que não é necessário que a pessoa tenha informação se estão ou não a ser adquiridas imagens. Há a possibilidade do observador poder rever as suas observações corrigindo erros que possam ter ocorrido. Evita também os principais problemas associados às técnicas de caracterização como o SOPLAY, em que o observador faz o varrimento da área de observação e depois tem um tempo de registo, tal faz com que exista perda de dados de registo. Com o uso de vídeos há a possibilidade de parar o vídeo e assim o observador realizar o registo ou até rever a sua observação.

Um dos problemas associados à utilização de vídeos está relacionado com a qualidade da imagem, à sua resolução e ao ângulo de visão. Na observação direta, utilizando o método

SOFIT, o observador necessita de observar continuamente a pessoa sem a perder do seu campo de visão. Com o recurso a vídeos é importante que exista essa preocupação, ou seja que a pessoa esteja sempre no campo de visão da câmara ou câmaras. Uma das soluções passa pelo uso de várias câmaras com diferentes ângulos de gravação que cubram toda a superfície ou área em análise.

O recurso a vídeos tem dois tipos principais de aplicações na caracterização da atividade física: podem ser aplicados para que um observador realize a caracterização ou usados para a classificação automática por aplicações computacionais.

Para uma eficiente utilização de vídeo para caracterização da atividade física é necessário, a existência de um estudo de metodologias, ferramentas e algoritmos que realizem a deteção do objeto ou objetos a serem caracterizados.

3. PROCESSAMENTO AUTOMÁTICO DE INFORMAÇÃO

Com crescente evolução da tecnologia associado ao processamento de vídeo, como a disponibilização de computadores de alto desempenho; câmaras de grande qualidade a baixo custo. Foi crescendo a necessidade para o desenvolvimento de análises automáticas de vídeo e para o desenvolvimento de algoritmos de seguimento para a caracterização do movimento humano. Tipicamente existem três passos cruciais neste tipo de análises: detecção dos objetos que interessam seguir; seguimentos dos mesmos objetos ao longo do tempo; interpretação dos resultados do seguimento do objeto para assim poder ser compreender o seu comportamento. O uso do seguimento de objetos é importante em tarefas como [25]:

- Reconhecimento baseado no movimento, isto é, identificação de humanos baseados nos seus movimentos;
- Sistemas de vigilância automáticos para monitorização e detecção de atividades suspeitas, ou eventos improváveis;
- Indexação de vídeo;
- Interação humano-computador como o reconhecimento de gestos e movimentos oculares;
- Monitorização de tráfego para possibilitar uma recolha de dados em tempo real do tráfego de automóvel.

3.1. SEGUIMENTO AUTOMÁTICO DE PESSOAS

O seguimento de objetos é definido por uma estimativa da trajetória de um objeto numa imagem onde ele se move durante uma sequência. Por outras palavras, seguimento consiste na atribuição consistente no tempo de uma etiqueta ou identidade (*label*) aos objetos. Adicionalmente, o algoritmo de seguimento pode retirar dados como a

orientação, área ou forma do objeto. O processamento de vídeo envolve a detecção de objetos e identificação das *frames* correspondentes de forma a existir correspondência entre os objetos e as frames seguintes. Isto pode ser entendido como a identificação de trajetórias dos objetos que se movem através do espaço observado [25].

Existem vários obstáculos à aplicação de algoritmos de seguimento como [25]:

- Movimentos da câmara;
- Ruído das imagens vídeo;
- Movimentos abruptos ou complexos dos objetos;
- Variação de iluminação;
- Ocultação parcial ou completa do objeto.

A segmentação dos objetos em vídeo, que consiste no agrupamento de regiões/pixéis numa imagem com base num determinado critério, é muitas vezes utilizada para a detecção dos objetos de interesse e sua trajetória. No entanto, existem múltiplos erros na segmentação de vídeo causados por sombras, mudanças de iluminação ou reflexões. Na Figura 3 (a), é apresentado uma imagem capturada por uma câmara de vigilância e a consequente máscara de segmentação *background-foreground* na Figura 3 (b), onde se verifica os problemas referidos devido a sombras e erros de detecção do limite do objeto [27].

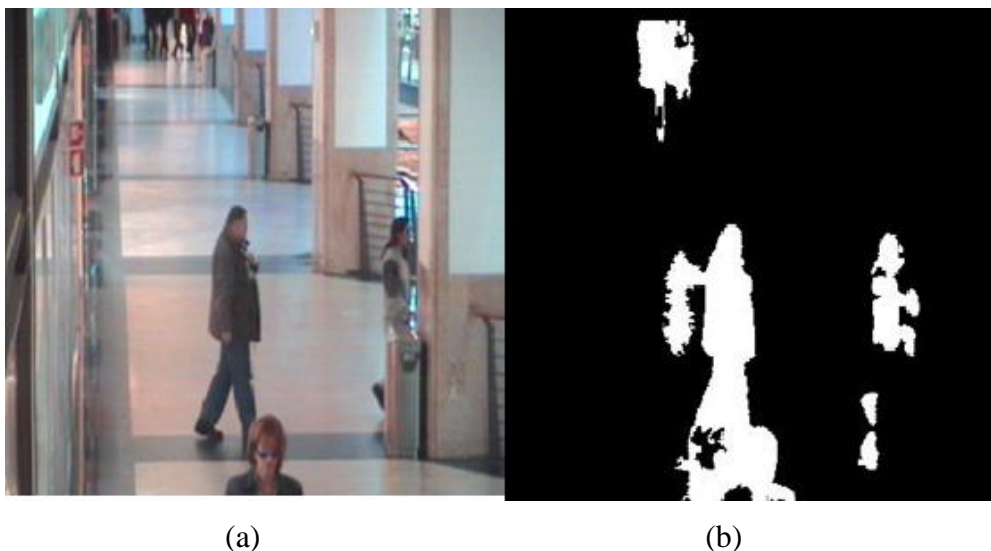


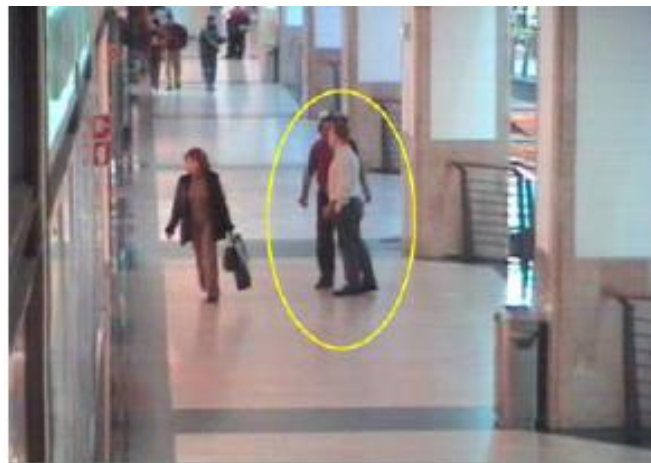
Figura 3 - Imagem capturada (a) e sua segmentação (b) do projeto CAVIAR [25].

Dois erros comuns que levam a uma segmentação errada são: divisão de um objeto em vários; Fusão de objetos num só. O primeiro caso ocorre quando um objeto é dividido em

vários objetos, o que vai criar um erro de classificação e vão ser seguidos independentemente. No segundo, vários objetos são seguidos como um só objeto (por exemplo, quando várias pessoas estão muito próximas), este tipo de problemas acontece no caso de ambientes com uma grande superfície de observação e a existência de grupos de pessoas. Um outro problema na segmentação de vídeo consiste na ocultação do objeto devido ao plano de fundo ou então por este se encontrar oculto por outro objeto, tal como se ilustram na Figura 4 (a), Figura 4 (b).



(a) – Ocultação devido background [25].



(b) - Ocultação devido a outro objeto [25].

Figura 4 – Diferentes tipos de ocultação [25].

Representação do Objeto

Um objeto de interesse é definido como algo relevante para uma futura análise, como por exemplo, barcos no oceano, pessoas na rua ou carros numa estrada [25]. Um objeto pode ser representado de diversas formas para a execução do seu seguimento [25]:

- Por pontos - O objeto é representado por um ponto, que se denomina centróide (Figura 5 (a)) ou por um conjunto de pontos (Figura 5 (b)). A representação por pontos geralmente é usada para o seguimento de objetos que ocupam uma pequena região da imagem;
- Formas geométricas - A forma do objeto é representado por uma figura geométrica como um retângulo ou uma elipse (Figura 5 (c), (d)). Este tipo de representação é usada para aproximar a forma humana e diminuir a complexidade do processamento;
- Representação pela silhueta ou contorno - A forma do objeto é representado pelos seus contornos (Figura 5 (g), (h)). A região dentro do contorno é chamado silhueta do objeto (Figura 5 (i)). Este tipo de representação é usado para seguimentos de objetos com formas não rígidas ou formas irregulares;
- Modelos de formas articuladas - Objetos articulados são compostos por partes do corpo humano que estão ligados entre si, como o torso, pernas, mãos, cabeça (Figura 5 (e)).
- Modelos baseado no esqueleto - É um modelo que é baseado no esqueleto do objeto (Figura 5 (f)).

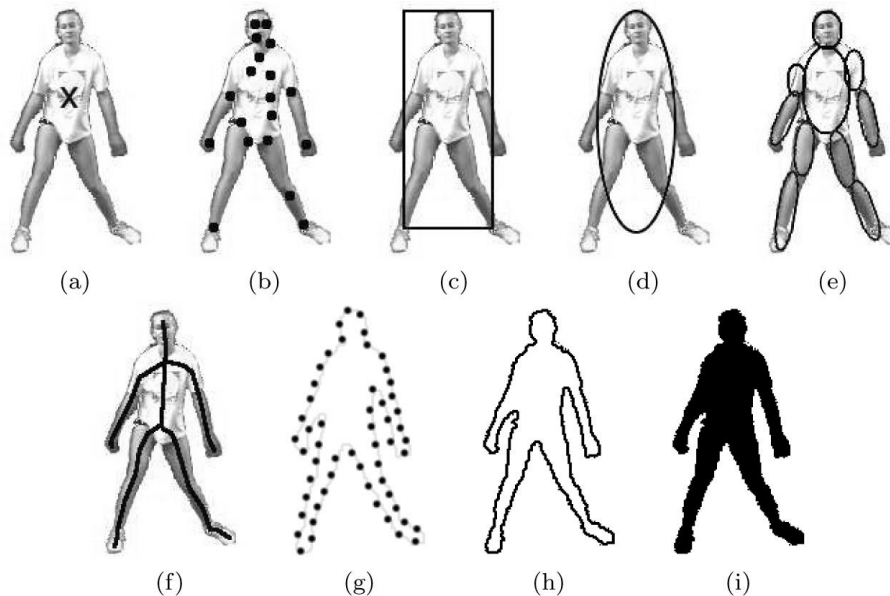


Figura 5 – Diferentes representações de um objeto para o seguimento [25].

No âmbito do uso de vídeos para o seguimento de pessoas existem diferentes estratégias que podem ser aplicadas. De uma forma geral, estas podem ser vistas de duas perspetivas diferentes [30]: *detection-based* e *matching-based*. A primeira baseia-se na existência de

uma correspondência temporal através do reconhecimento de objetos em *frames* diferentes, enquanto na segunda é estimada a trajetória de cada objeto.

Independentemente do funcionamento de cada algoritmo, tipicamente existem três aspectos comuns: a identificação dos aspectos revelantes numa sequência, como por exemplo a identificação de pessoas a serem seguidas, contendo também outros objetos em movimento. Uma vez detetados, os objetos são transformados para uma representação diferente, tornando mais fácil o seu processamento pelos algoritmos [28][31][32]. Por fim os objetos são seguidos *frame-a-frame*. Uma das estratégias mais comuns é *predict-match-update* que é designado previsão-correspondência-atualização, é ilustrado nas figuras abaixo. A abordagem consiste na previsão de uma posição x a um instante de tempo t , com base nas informações existentes num instante tempo $t-1$ (linha amarela na Figura 5 (a)), isto é, para um instante de tempo t é realizado uma procura nas proximidades de uma posição x pelo objeto (Figura 5 (b)). Se o objeto for encontrado a informação é atualizada com a nova observação (Figura 5 (c)) [33].



(a) – Ilustração do método *predict* [25].



(b) – Ilustração do método *match* [25].



(c) – Ilustração do método *update* [25].

Figura 6 – Ilustração do método *predict-match-update* [25].

No seguimento de objetos, a seleção das características que possam ajudar ao seguimento é de uma elevada importância. Em geral são selecionadas características únicas ao objeto que permitam identificar o objeto de forma inequívoca. A seleção de características está diretamente relacionada com o tipo de seguimento e sobretudo com o tipo de objeto que se pretende seguir [26][25]. Existem várias características frequentemente utilizadas na literatura como:

- Cor.
- Limites do objeto.
- Textura;
- Forma.

3.2. MACHINE LEARNING E DATA MINING

Técnicas de máquina de aprendizagem (*Machine Learning*) tem sido bastante utilizadas na literatura em tarefas com previsão baseada em dados passados ou classificação automática. *Machine Learning* é uma área de inteligência artificial que foca no desenvolvimento de programas capazes de aprender e de se adaptarem de acordo com os dados que são fornecidos. Esta aprendizagem utiliza uma quantidade de observações ou dados como exemplos de uma situação. O grande objetivo de *Machine Learning* é a criação de métodos automáticos, ou seja é a criação de algoritmos eficazes e inteligentes o suficiente para a realização de uma tarefa sem a interação ou assistência humana. *Machine Learning* pode ser comparado, como quando o ser humano é confrontado com situações e tarefas diárias e para as resolver usa ferramentas específicas que já foram usadas anteriormente [35].

Alguns exemplos de problemas associados a *Machine Learning* são:

- Detecção facial;
- Detecção de *spam*: como em emails, a distinção de emails de *spam* e *non-spam*;
- Categorização de tópicos: notícias, desporto, política, entretenimento, etc.;
- Detecção humana e previsão dos seus movimentos;
- Previsão meteorológica.

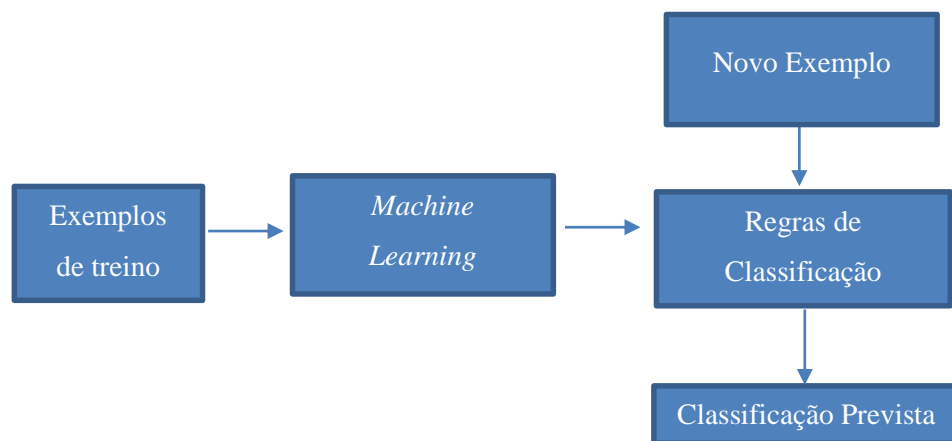


Figura 7 – Diagrama típico de um problema de aprendizagem (adaptado de [35]).

Data Mining consiste na funcionalidade que agrega e organiza dados, encontrando assim padrões, associações, mudanças e anomalias relevantes. É uma técnica que utiliza várias metodologias, como a inteligência artificial, reconhecimento de padrões e estatística para procurar correlações entre diferentes dados.

Data Mining é:

- Exploração;
- Construção do modelo;
- Definição do padrão;
- Validação;
- Verificação.

O objetivo é encontrar relações, similaridades entre os movimentos humanos para assim poder determinar as interações entre eles.

Construção de um Modelo de Classificação

Para realizar uma classificação através de *Machine Learning*, é necessário criar um modelo de classificação. É necessário a criação de classes de classificação, como por exemplo: sedentário, moderado, vigoroso, e todos os elementos que possam contribuir para diferenciar cada classe.

No processo de classificação a cada classe é atribuído uma *label* de acordo com as suas características, posteriormente é usado um classificador que usa todas as características anteriormente dadas e atribui uma das diferentes classes definidas.

Propriedades usadas para representação de classes

Como referido anteriormente, para a criação de um classificador é necessário a existência prévia de elementos que possam ajudar na classificação de um objeto.

Por exemplo, no caso de movimento de pessoas em vídeos, elementos como:

- Posição que se encontra na imagem;
- Área ocupada;
- Tamanho da caixa (largura/altura);
- Orientação da pessoa.

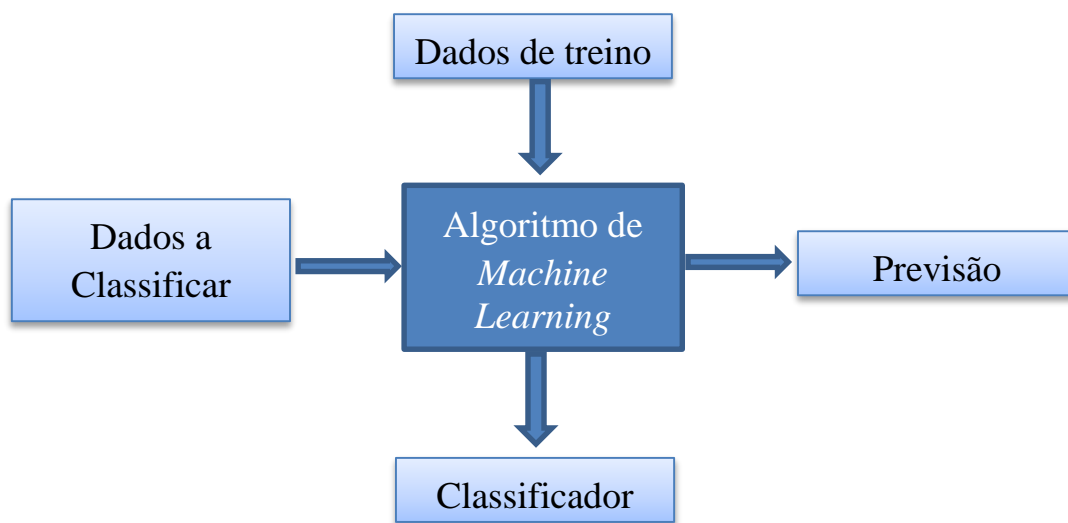


Figura 8 – Modelo de classificação

Na Figura 8 está representado o procedimento de criação do classificador até a previsão de novos dados. Inicialmente é necessário a criação de um classificador, tal pode ser realizado através de dados de *input* que contém todas as classes e elementos que as constituem. Depois de treinado o classificador, este pode ser utilizado para “etiquetar” a informação de entrada das possíveis classes. Existem diversas implementações de algoritmos de *machine learning*, destacando-se as bibliotecas de LIBSVM, WEKA, MATLAB.

Existe diferentes métodos de apoio à decisão na classificação estes definem-se por todas as metodologias que ajudam na construção do modelo para a classificação. São técnicas que ajudam na decisão quais os elementos que possibilitam uma classificação de uma

determinada classe com mais exatidão ou auxiliam na análise de resultados obtido dos diferentes classificadores.

Métodos tais como:

- Matrizes de confusão: são utilizadas para ver o resultado e avaliar o output;
- Árvores de decisão: são um mecanismo de classificação/decisão;
- *Cross-Validation*: processo usado para validação de um classificador.

Árvores de decisão

Árvores de decisão são uma estrutura em forma de árvore que tem uma função de decisão em cada nó, cada decisão em cada nó vai determinar a classificação final. As árvores de decisão são utilizadas melhor entendimento das características de cada classe e o que as diferencia de outra classe. A árvore de decisão é uma estrutura compacta que mostra uma possibilidade de classificação executada pelo classificador.

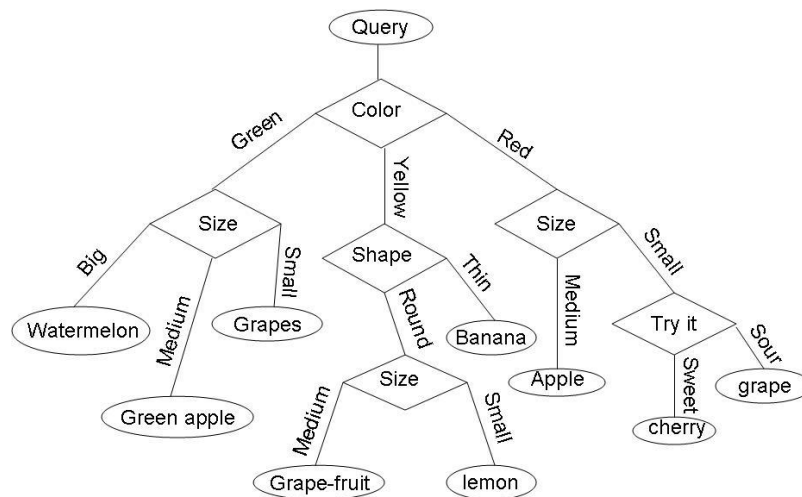


Figura 9 – Exemplo Árvore de decisão [36].

Na Figura 9 é exemplificado uma possível árvore de decisão e uma possível classificação.

Esta árvore demonstra um classificador de frutos, onde é possível ver as diferentes classes como *watermelon/green apple/grapes* etc., como as suas características. O processo de classificação começa desde a cor até ao seu tamanho ou sabor.

Support Vector Machine

Os *Support Vector Machine* (SVM) são um tipo de modelos de aprendizagem associados a algoritmos que analisam dados e reconhecem padrões nesses dados. Dado um número de exemplos de treino todos marcados como pertencentes a diferentes classes, o algoritmo SVM constrói um modelo que atribui novos exemplos as classes existentes [38][41].

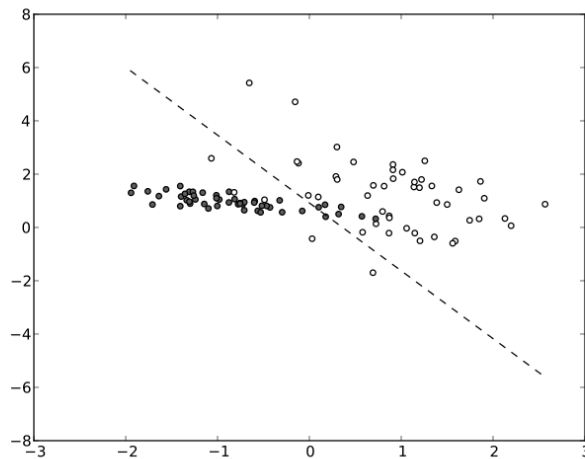


Figura 10 – Possível mapeamento de diferentes classes pelo SVM em uma classificação Linear [44].

Na Figura 10 pode ser verificado um exemplo do mapeamento de diferentes classes pelo SVM, os pontos correspondem a uma classe e são diferenciados por classes diferentes, o que o SVM realiza é novos exemplos são mapeados no mesmo espaço de maneira a que seja possível verificar qual a categoria a que pertencem, baseado no local onde são mapeados.

SVM além de fazer uma classificação linear também é capaz de executar numa classificação não linear usando um *kernel*, implicitamente mapeando todos os *input's* em espaços multidimensionais. Originalmente o SVM foi desenvolvido para trabalhar com classificadores lineares mas com a evolução deste tipo de sistemas foi necessário a criação de classificadores não lineares e foi então que surgiu o uso de um *kernel* [41].

O truque do uso de um *kernel* é simplesmente tornar um espaço 2D em 3D, isto é numa classificação linear é possível separar cada classe por espaço (Figura 10). O *kernel* permite encontrar um plano onde seja possível separar as duas classes, por exemplo num caso com duas classes o SVM encontra um plano onde as possa separar. No caso da Figura

11, é impossível existir uma separação linear das duas classes, ou seja o SVM não encontrar uma solução para separação das classes, mas com o uso de um *kernel* é possível encontrar uma solução [39].

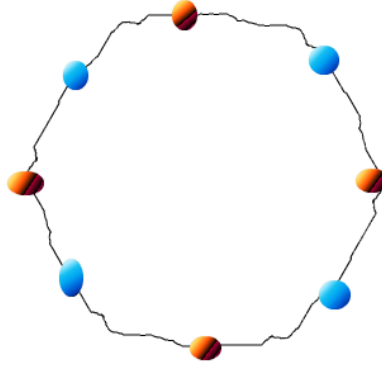


Figura 11 – Exemplo duas classes onde é impossível o SVM separá-las [40].

O *kernel* ajuda neste tipo de situações e encontra um plano que possa separar as duas classes através de algoritmos não lineares sem comprometer as propriedades das duas classes [40].

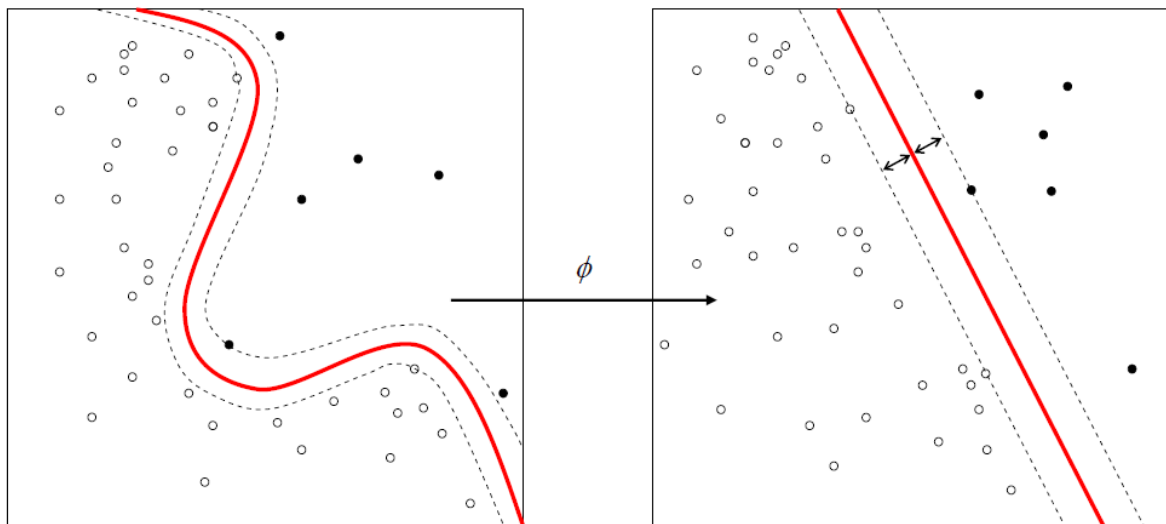


Figura 12 – Exemplo da transformação realizada por um *kernel* [42].

Matrizes de confusão

São normalmente usadas para retirar relatórios dos resultados dos testes de classificação e na verificação da performance do método. Estas matrizes permitem verificar quais as classes que estão com maior acerto e onde existe maior taxa de classificação errada e maior confusão em cada classe.

Tabela 1 – Exemplo de uma tabela de confusão

	Classe j	“0”	“1”	“2”	“3”	“4”	“5”	“6”
Classe i	“0”	99	0	0	0	0	0	1
	“1”	0	98	0	0	2	0	0
	“2”	0	0	96	1	0	3	0
	“3”	0	0	2	95	0	5	0
	“4”	0	0	0	0	98	2	0
	“5”	0	0	0	2	0	98	0
	“6”	1	0	0	0	0	1	98

Na Tabela 1 apresenta-se um exemplo de uma tabela de confusão de classificação de dígitos. É possível verificar o número de classificações erradas e corretas de cada dígito. Com este método é possível verificar se os elementos que definem cada classe estão corretos através da taxa de acerto na classificação de cada dígito. A diagonal onde $i=j$, indica a taxa de sucesso, todos os outros elementos fora dessa diagonal que não sejam iguais a zero indicam que existiu confusão com esses dígitos, ou seja se a taxa desses elementos for alta significa que os elementos extraídos para cada classe não estão totalmente corretos, estão a criar confusão na classificação.

Cross-Validation

A validação cruzada (*Cross-Validation*) é uma técnica muito utilizada na avaliação de classificadores, sobretudo quando os dados disponíveis são limitados. Com esta técnica é possível generalizar um conjunto de dados através de uma análise estatística. Aquando da criação de um classificador são usado um conjunto de dados de treino onde é conhecido todas as classes e suas características para suas previsões (*dataset* de treino) e um outro de conjunto de dados desconhecidos (*dataset* de teste). O objetivo com do uso de *cross-validation* é realizar uma validação do classificador, com um conjunto de dados estatísticos e verificar o *output* do classificador.

3.3. BIBLIOTECAS COM SUPORTE PARA *MACHINE LEARNING*

3.3.1. WAIKATO ENVIROMENT FOR KNOWLEDGE ANALYSIS - WEKA

Weka (*Waikato Enviroment for Knowledge Analysis*) é uma aplicação *Open Source*, foi desenvolvida na Universidade de Waikato na Nova Zelândia. É uma aplicação grátis, desenvolvida usando Java. Weka tem como objetivo agregar algoritmos provenientes de diferentes abordagens/paradigmas na área da inteligência artificial dedicada a *Machine Learning*, e procede à análise computacional e estatística dos dados recorrendo a técnicas de *data mining* tentando a partir de padrões encontrados, pesquisar uma solução. Weka contém módulos de processamento de dados, classificação, associação, regra de associação e uma interface gráfica, do qual permite escolher alguns detalhes e agilizar o processo de criação de *Machine Learning*.

As principais características do Weka são:

- Diferentes ferramentas de processamento de dados;
- Diferentes algoritmos de classificação;
- Algoritmos de associação (*clustering*);
- Avaliação de atributos,
- Algoritmos para encontrar regras de associação;
- Diferentes interfaces para o utilizador
 - “The Explorer” (análise exploratória de dados, pré-processamento, seleção de atributos, aprendizagem e visualização);
 - “The Experimenter” (Ambiente experimental, teste e avaliação de algoritmos de *Machine Learning*);
 - “The KnowledgeFlow” (Ambiente de visualização dos processos KDD).
- Linha de comandos simples.

3.3.2. MATRIX LABORATORY – MATLAB

MATLAB é um *software* que é usado para o cálculo numérico, usa uma linguagem de que integra computação, visualização e ambiente de programação. MATLAB usa uma linguagem de programação moderna e fácil para a escrita de programas para programação matemática, pois contém estruturas de dados sofisticadas e ferramentas de *debugging*, criação de gráficos, produzindo figuras em 2D e 3D e as editar e suporte de programação

orientada a objetos. Estes são alguns dos fatores que tornam o MATLAB uma ferramenta muito usada em investigação e no ensino.

O *software* MATLAB contém poderosas rotinas que possibilitam uma vasta variedade de computações. É uma ferramenta de fácil uso pois disponibiliza comandos gráficos que permitam a visualização imediata de resultados. Também tem diferentes ferramentas para processamento de sinal, teoria de controle, simulação, otimização. Possui um extenso suporte de álgebra linear e operações com matrizes e várias outras ferramentas que podem ser aplicadas no âmbito de engenharia.

3.3.3. BIBLIOTECAS MACHINE LEARNING NO PYTHON

Python é uma linguagem de programação de alto nível orientada a objetos foi criada por Guido van Rossum, é uma linguagem de alto nível isto porque é construído por estruturas de dados, combinado com ligações dinâmicas. Python é uma linguagem que foi desenhada para ser simples, fácil de aprender e de grande portabilidade [44].

Existem varias bibliotecas de suporte de *machine learning* para o Python, tais como o Monty e o Elephant. São bibliotecas que incluem diferentes algoritmos de otimização e resolução de problemas de *Machine Learning*, são *open source* e incluem uma interface de utilizador para uma rápida criação de um protótipo de *Machine Learning*. Em *machine learning* é comum o uso da linguagem Python devido à sua grande capacidade e fácil utilização. Tem um conjunto de ferramentas desenvolvidas para o uso nas diferentes bibliotecas de *Machine Learning* [45][46].

3.3.4. OPENCV

OpenCV é uma biblioteca multiplataforma livre usada para o desenvolvimento de aplicações na área de visão computacional e *machine learning*, possui módulos de processamento de imagens e vídeo, estrutura dados, álgebra linear e algoritmos de visão computacional, como reconhecimento de objetos, filtros de imagem, classificação de movimentos humanos em vídeos. É baseada nas linguagens de C/C++ e possui suporte para linguagens como JAVA, Python, Visual Basic [50][51].

OpenCV foi lançado em 1999 pela Intel Research com o objetivo de melhorar aplicações de processamento no campo de visão computacional. O principal objetivo era promover a pesquisa em visão computacional e não apenas o código aberto, mas também otimizado

de forma a pode ser facilmente lido e transferível. A sua estrutura contém vários módulos como [50]:

- Cv – módulo de funcionalidades e algoritmos de Visão computacional;
- Cvaux – módulo com algoritmos de visão;
- ML – módulo de *machine learning*, suportando algoritmos de classificação, árvores de decisão, *k-nearest neighbors*, etc.;
- Etc.

3.3.5. **LINGUAGEM R**

R é uma linguagem de *software* e ambiente gráfico de análise estatística comum entre os utilizadores de análise estatística e de *data mining*. R dispõe de uma variedade de técnicas de estatística e de gráficos, incluindo modelos lineares e não lineares, testes clássicos estatísticos, classificação. A linguagem R provem de uma outra linguagem, a linguagem S, faz parte da GNU *project* e o código fonte foi criado num ambiente escrito em C. Usa a linha de comandos como sua interface, mas já existe várias GUI para o seu uso. Linguagem R, tal como outras linguagens similares (MATLAB, APL), suporta matrizes aritméticas, estruturas contem vetores e matrizes, *arrays*, *data frames* e listas. A linguagem R suporta algoritmos de *Machine Learning*, porque é uma linguagem que todos os seus módulos aplica testes estatísticos, contem módulos gráficos e funções matemáticas.

3.3.6. **LIBSVM**

LIBSVM é uma biblioteca integrada no SVM (*support vector machine*), que usa algoritmos com *kernel's* para classificação e regressão linear ou não linear. O seu desenvolvimento teve como objetivo facilitar aos utilizadores a implementação de SVM nas suas aplicações. O LIBSVM é das bibliotecas com maior popularidade em sistemas que usam *Machine Learning* [52] e suporta classificação com múltiplas classes. Na Figura 13 e Figura 14 visualiza-se um exemplo típico de um output do uso do LIBSVM, este envolve normalmente duas etapas: uma primeira é de treino de um conjunto de dados para obter o modelo de treino.


```

c:\libsvm-3.0\windows>svm-train.exe -c 2.0 -g 0.5 data.1.scale
*
optimization finished, #iter = 21
nu = 0.311375
obj = -22.038870, rho = -0.237857
nSV = 16, nBSV = 12
.*
optimization finished, #iter = 93
nu = 0.289157
obj = -32.831806, rho = 0.885004
nSV = 25, nBSV = 22
*
optimization finished, #iter = 41
nu = 0.103896
obj = -9.039803, rho = 0.544485
nSV = 9, nBSV = 6
Total nSV = 39
c:\libsvm-3.0\windows>

```

Figura 13 – Exemplo de criação modelo de treino.

Uma segunda etapa é envolver o uso do mesmo modelo para a classificação de um conjunto de dados de teste, do qual resulta um conjunto de dados de saída que contém a previsão desse mesmo conjunto de dados de teste e a sua percentagem de precisão na classificação. Esta biblioteca, inclui ferramentas de apoio para o uso de *kernel* onde é ser possível a otimização do *kernel* para o modelo de treino ser o mais preciso possível.

```

c:\libsvm-3.0\windows>svm-predict.exe data.1.scale data.1.scale.model data.1.out
Accuracy = 99.0476% (104/105) (classification)
c:\libsvm-3.0\windows>

```

Figura 14 – Exemplo de classificação e percentagem de precisão do modelo

4. PREPARAÇÃO E CARACTERIZAÇÃO DOS DADOS

Com necessidade de comprovar a utilidade e melhoria que a introdução de vídeos pode trazer na observação direta foi necessário a realização de diferentes experiências com diferentes observadores e realizar uma análise sobre os dados que podiam ser extraídos das sequências vídeo. Foi necessário a realização de diferentes experiências com vídeos de maneira a demonstrar que os observadores teriam menos stress na observação, conseqüentemente teriam melhor qualidade na visualização, anotações mais precisas, porque poderiam andar para trás se encontrassem dúvidas.

Foram realizadas duas experiências e uma geração de dados. Foi realizada uma experiência onde se pretendia simular a observação direta tradicional, sem pausas nos vídeos e uma segunda experiência em que haveria pausa nas sequências de maneira que os observadores poderiam anotar e assim não existir uma perda de dados. A geração de dados foi necessária para existir um termo de comparação nas anotações e ser possível verificar seu nível de precisão, suas limitações e vantagens de cada experiência. Foi realizada com os três observadores de maneira a existir uma total concordância entre os três, houve um período de discussão de maneira a tornar cada anotação a mais correta. O resultado desta fase será tomado como informação de referência.

4.1. CONJUNTO DE DADOS

Para a análise da utilização de vídeos na caracterização da atividade física foram capturadas, no decorrer desta dissertação, sequências vídeo com diferentes pessoas a realizarem diversos movimentos (como por exemplo correr, saltar, andar). Foram observadas 10 pessoas divididas em dois grupos de 5 elementos que desempenharam as diversas atividades em simultâneo. Os grupos ocuparam duas áreas distintas do espaço, cada uma coberta por uma câmara. Esta disposição dos sensores e divisão das pessoas

permite recriar cenários onde: um observador consiga cobrir toda a área; onde são necessários 2 observadores ou a observação de cada grupo tem de ser efetuada em instantes disjuntos. As câmaras foram colocadas em posições típicas de monitorização: numa posição elevada, inclinada para baixo, de modo a aumentar a área coberta e diminuir as situações de ocultação.



Figura 15 - Perspetiva da sequência de um grupo.



Figura 16 – Perspetiva da sequência de um grupo.

Foram utilizadas 2 câmaras IP, (Figura 15 e Figura 16), de alta definição diferentes: Sony SNC-CH120; Vivotek FD8162. Os vídeos foram capturados com características semelhantes que se encontram resumidas na Tabela 2.

Tabela 2 – Propriedades das sequências de vídeo

Duração (minutos)	Resolução	Frame/Rate	Codec	Nº Pessoas
8:40	1024x768	30 frames/s	mp4	5

4.2. DESCRIÇÃO METODOLOGIAS DE CARACTERIZAÇÃO E ANOTAÇÃO

Neste trabalho pretende-se analisar e caracterizar os métodos de observação direta, sob o ponto de vista dos observadores, e de forma a informação visual sob a forma de vídeo pode ser utilizada para maior fiabilidade dos resultados. Assim, e sendo o SOPLAY visto como referência em observação direta, esse método foi escolhido para caracterização.

A análise das sequências capturadas foi realizada por 3 observadores, devidamente treinados neste método, e em 2 experiências:

1. Anotação sem pausas do vídeo – corresponde ao método SOPLAY original, onde um período de observação é seguido de um período de anotação, neste caso existe perda de informação correspondente ao tempo em que o observador anota o que visualizou;
2. Anotações com pausas – a diferença relativamente ao método original consiste na paragem do vídeo durante o período em que é feita a anotação, desta forma não existe perda de informação;

Foram realizadas também anotações com pausas e possibilidade de nova visualização – nesta anotação é possível voltar atrás no vídeo para resolver ambiguidades e garantir precisão, para assim ser possível existir um modelo de referência para comparar os dados das experiências com um modelo.

No sentido de não existir interferência nas observações realizadas, estas foram realizadas por cada observador individualmente. Cada fase foi realizada com o espaçamento de duas ou mais semanas de forma a evitar enviesamentos ou fadiga dos observadores.

Para a realização das observações foi necessário a criação de um formulário que preenchesse e ajudasse os observadores a realizarem uma observação o mais correta possível e próxima da observação direta sem o uso de vídeos.

Este formulário foi distribuído aos observadores e continha:

- Indicação do vídeo a observar;
- Data;
- O tempo de início do vídeo e tempo que havia paragem para anotação;
- Tipo de ambiente (interior ou exterior);
- Número de Observações;
- Os diferentes tipos de caracterização (Sedentário/Andar/Vigoroso);
- Grau de confiança do observador em cada observação.

Observador:		Video:					
				Laranjas			
start date	start time	stop time	Scan	Sed	Wal	Vig	Confiança
14/04/2014	00:00:00	00:00:10	1				
	00:00:10	00:00:20	2				
	00:00:20	00:00:30	3				
	00:00:30	00:00:40	4				
	00:00:40	00:00:50	5				
	00:00:50	00:01:00	6				
	00:01:00	00:01:10	7				
	00:01:10	00:01:20	8				
	00:01:20	00:01:30	9				
	00:01:30	00:01:40	10				
	00:01:40	00:01:50	11				
	00:01:50	00:02:00	12				
	00:02:00	00:02:10	13				

Figura 17 – Excerto formulário de anotação

A Figura 17 contém um excerto de um formulário de anotação. Os formulários são todos semelhantes para todas as fases de anotação realizadas, a única diferença foi em relação ao *Start Time* e *Stop Time*. Nas anotações com paragens, não há perda de informação e todo o vídeo é caracterizado. Consequentemente o número de observações vai ser maior de que na fase onde não existe paragem para anotação, pois nesta última o tempo é contínuo e o número de anotações vai descer para metade isto porque há perda de informação (i.e., idêntico à observação direta tradicional).

Devido à subjetividade que a análise por um observador implica, foi necessário incluir no formulário um campo que capturasse o nível de confiança do observador em cada observação.

4.3. CARACTERIZAÇÃO DOS DADOS DA OBSERVAÇÃO DIRETA

Nesta fase foi realizado uma caracterização da observação direta através do ponto de vista do observador. Procurou-se analisar a precisão de cada observação e o efeito que terá a subjetividade desta caracterização. Para cada fase de observação foram realizadas diferentes medidas:

- Precisão: número de classificações corretas quando comparado com a informação de referência;
- Precisão com concordância: número de classificações corretas (de acordo com a informação de referência) e com concordância com pelo menos mais um observador;
- Taxa de erro: número de classificações erradas e com desacordo entre os observadores;
- Taxa de erro com concordância: número de classificações onde pelo menos dois observadores concordam entre si apesar de estarem errados em relação à informação de referência.

Após a conclusão das diferentes experiências foi realizada uma caracterização dos dados. No caso da primeira experiência onde não há paragem dos vídeos, (52 observações/vídeo), opõem-se a comparação dos dados dos 3 observadores com a anotação de referência, (Tabela 3 e Figura 18).

Tabela 3 – Caracterização dados Anotação sem paragens

	Observador 1 (%)	Observador 2 (%)	Observador 3 (%)	Media (%)
Precisão	50.00	57.69	36.54	50.00
Precisão com concordância	38.46	44.23	34.62	38.46
Taxa de erro	40.38	25.00	51.92	40.38
Taxa de erro com concordância	9.62	17.31	11.54	11.54

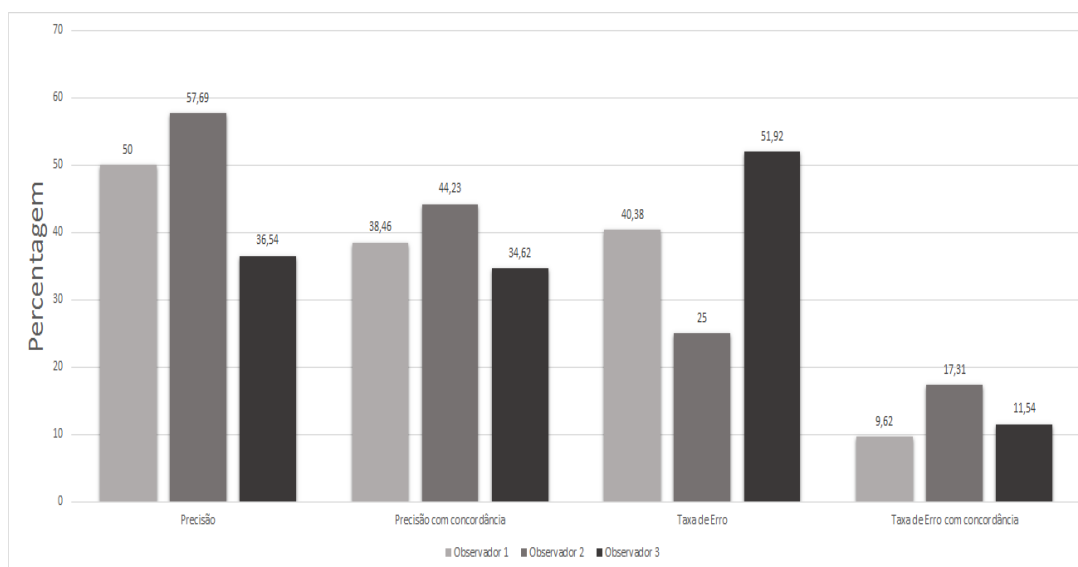


Figura 18 – Anotações sem paragens no conjunto do dos vídeos

Após a análise dos resultados, verifica-se que os observadores de uma maneira geral acertam em pelo menos metade das anotações, quer em concordância com o modelo de referência, como em concordância com outro observador.

Na experiência 2, como há paragens nos vídeos para anotação, não há perda de informação, toda a sequência é anotada, (104 observações no total), os dados foram os seguintes:

Tabela 4 - Caracterização dados Anotação com paragens

	Observador 1 (%)	Observador 2 (%)	Observador 3 (%)	Media (%)
Precisão	55,77	40.38	42.31	42.31
Precisão com concordância	41.35	36.54	33.65	36.54
Taxa de erro	27.88	41.35	46.15	41.35
Taxa de erro com concordância	16.35	18.27	11.54	16.35

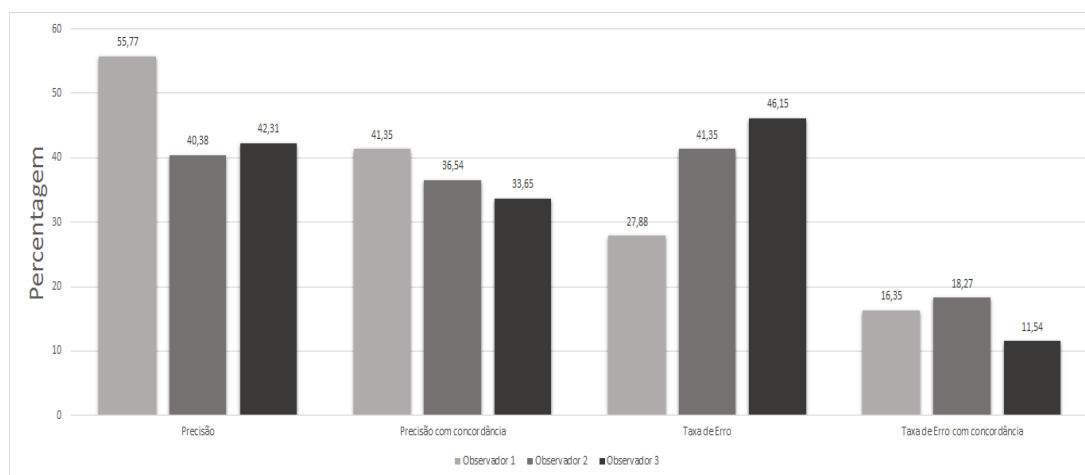


Figura 19 - Anotações com paragens no conjunto dos vídeos

Na Tabela 4 e Figura 19 a grande diferença verifica-se na taxa de erro, pois esta desceu substancialmente e aumentou a precisão em dois dos três observadores. O que significa, que com paragens os observadores não tem a pressão de anotar o mais rápido possível devido ao tempo de anotação, o que aumenta a confiança e leva ao aumento do acerto, de uma maneira geral, quer com o modelo de referência como entre eles.

4.3.1. COMPARAÇÃO DAS DUAS EXPERIÊNCIAS DE ANOTAÇÃO

No âmbito da verificação das grandes diferenças entre as duas experiências de anotação foi necessário realizar uma comparação de dados com os mesmos intervalos de tempo, para isso foi necessário retirar todos os dados em excesso da experiência 2 em relação à experiência 1. Na experiência 2 não tem paragens, não há perda de informação logo, vai existir o dobro das observações, sendo assim retira-se os dados em excesso, que correspondem aos intervalos de tempo que na experiência 1 os observadores estão a anotar e não a observar.

Tabela 5 – Dados comparativos das duas fases de anotação

	Observador 1 (%)		Observador 2 (%)		Observador 3 (%)		Media (%)	
	Exp. 1	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2
Precisão	50,00	53,85	57,69	44,23	36,54	46,15	48,08	48,08
Precisão com concordância	38,46	42,31	44,23	38,46	34,62	38,46	39,10	39,74
Taxa de erro	40,38	32,69	25,00	36,54	51,92	40,38	39,10	36,54
Taxa de erro com concordância	9,62	13,46	17,31	19,23	11,54	13,46	12,82	15,38

Na comparação entre as duas experiências era de esperar que na experiência 2 a precisão aumentasse e a taxa de erro diminuísse. Na Tabela 5 podemos verificar que existiu uma ligeira melhoria nos resultados, a precisão teve um ligeiro aumento e a taxa de erro diminuiu. Um outro dado importante é a existência de uma melhoria da concordância entre observadores e com o modelo de referência. Podemos concluir que os resultados da experiência 2 são melhores, mas essa melhoria foi muito ligeira, o que demonstra a subjetividade da observação direta devido ao uso de observadores, o que leva a concluir que com uma classificação automática poderá existir uma grande margem de melhoria na precisão, como da objetividade.





Em relação aos tempos de cada tipo de anotação (Tabela 6), como era de esperar, o tempo foi aumentando em cada experiência. Onde o tempo da anotação sem paragens foi tomado como referência, isto porque é o tempo da sequência de vídeo. Como era de prever, a criação do modelo de referência foi a anotação mais prolongada, porque além da existência de paragens também existia discussão entre os observadores para chegarem a um consenso.



Tabela 6 – Tempo realizado para cada tipo de anotação.

Tempos de observação			Tempo Total (minutos)	Aumento do tempo de observação (%)
	Vídeo 1 (minutos)	Vídeo 2 (minutos)		
Sem pausas	8,4	8,4	16,8	-
Com Pausas	19,33	18,15	37,48	123,1
Observação referência	37,12	27,03	64,15	281,8

Seguidamente foi realizada uma análise para procurar identificar onde existe maior confusão na classificação e consequentemente à falha no tipo de movimento. Para isso foi realizado um cruzamento de dados entre os observadores onde se poderia identificar quais os movimentos em que cada observador cria mais confusão.

Legenda:

-  - observador acertou no tipo de movimento do observado;
-  - foi classificado erradamente como movimento sedentário;
-  - foi classificado erradamente como movimento moderado;
-  - foi classificado erradamente como movimento vigoroso;

-  - dividido por dois tipos de movimento, observador errou em mais que um tipo de movimento em uma anotação;
-  - observador classificou erradamente o numero de pessoas presentes.

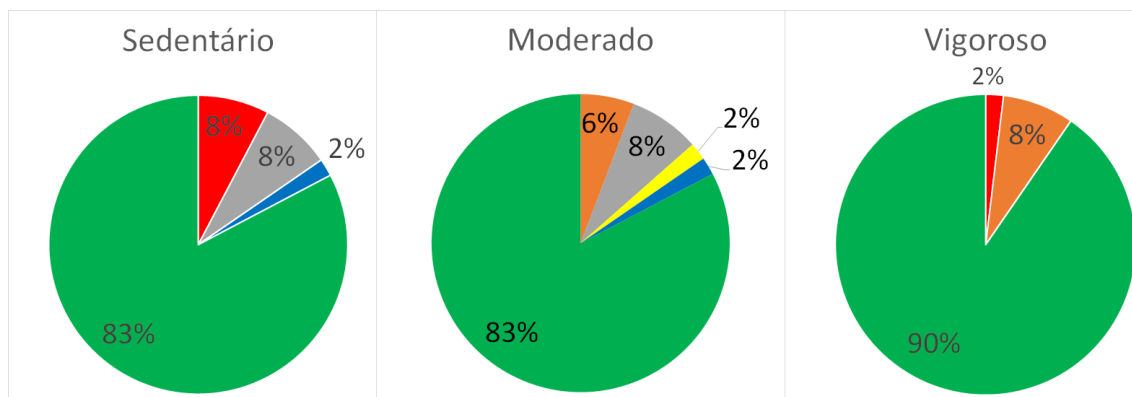


Figura 20 – Gráfico de confusão por tipo de movimento para o Observador 1 sem paragens

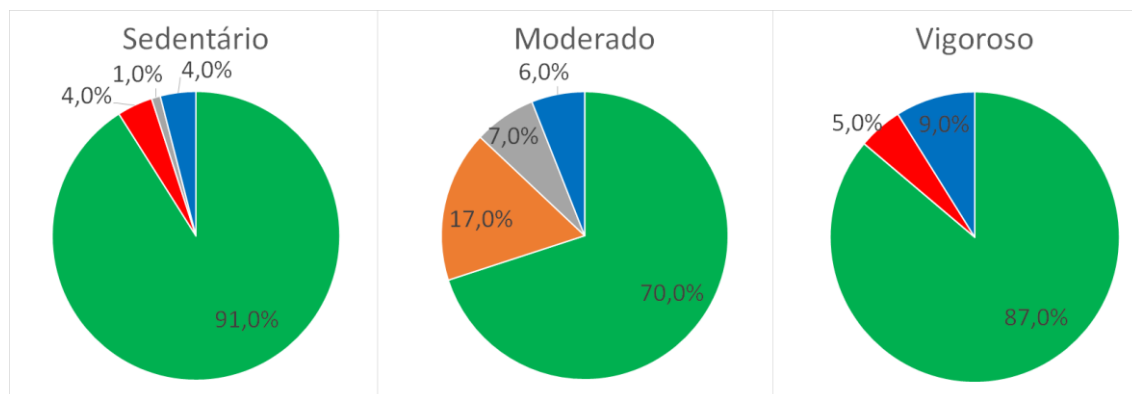


Figura 21 – Gráfico de confusão por tipo de movimento para o Observador 1 com paragens

Nas Figura 20 e Figura 21 é possível verificar que para o movimento sedentário e moderado o observador 1 tem uma taxa de erro igual nas duas experiencias. Nesses mesmos movimentos tem a taxa de acerto elevada contudo existe uma confusão com os outros tipos de movimentos e uma pequena taxa de falha na contagem das pessoas intervenientes. No movimento vigoroso, como era de esperar existe uma confusão com a classe mais próxima (moderado).

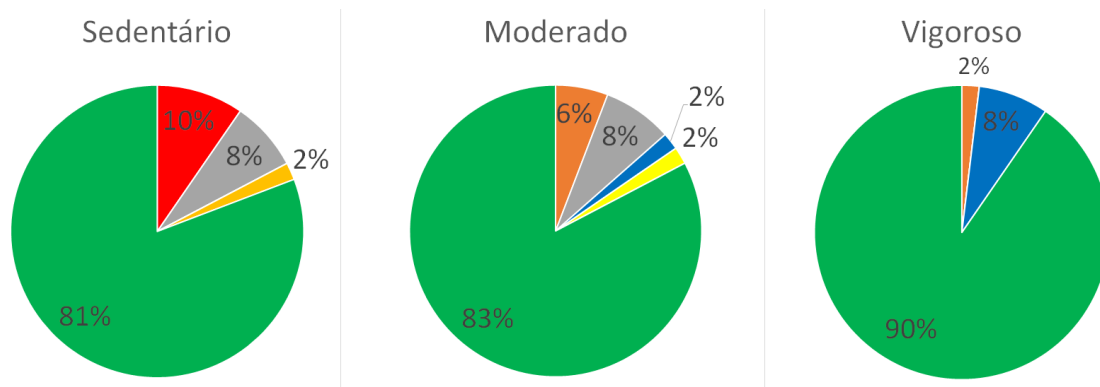


Figura 22 – Gráfico de confusão por tipo de movimento para o Observador 2 sem paragens

Na Figura 22, ao contrário do esperado, no movimento vigoroso existe uma pequena confusão com o movimento sedentário, esta devesse concluir ao curto tempo de anotação que o observador dispõe, o que levou a um erro de anotação. Nos movimentos, sedentário e moderado, os resultados são semelhantes aos do observador 1. Já as Figura 23, Figura 24 e Figura 25 vão de encontro ao que era esperado (cada movimento com confusão com os movimentos mais parecidos), visto que o observador confunde, em cada movimento com o movimento mais semelhante, ou seja o sedentário o movimento moderado e no movimento vigoroso maior confusão com o movimento moderado.

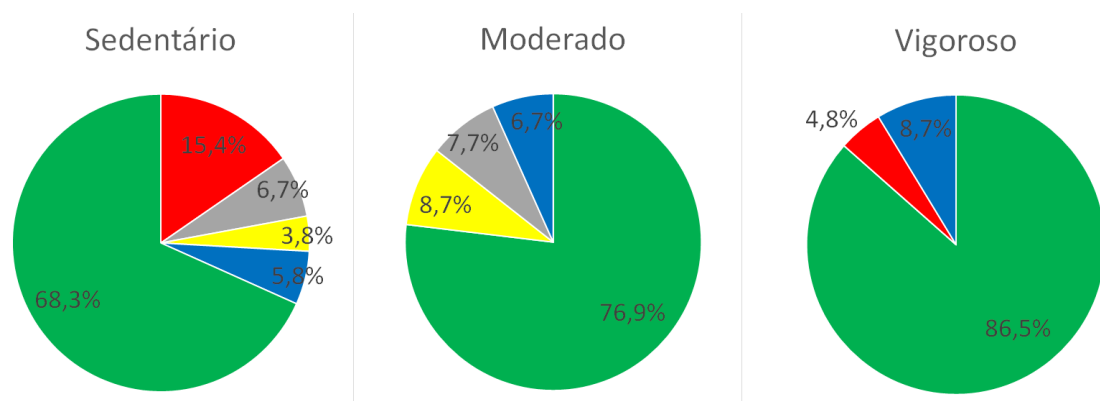


Figura 23 - Gráfico de confusão por tipo de movimento para o Observador 2 com paragens

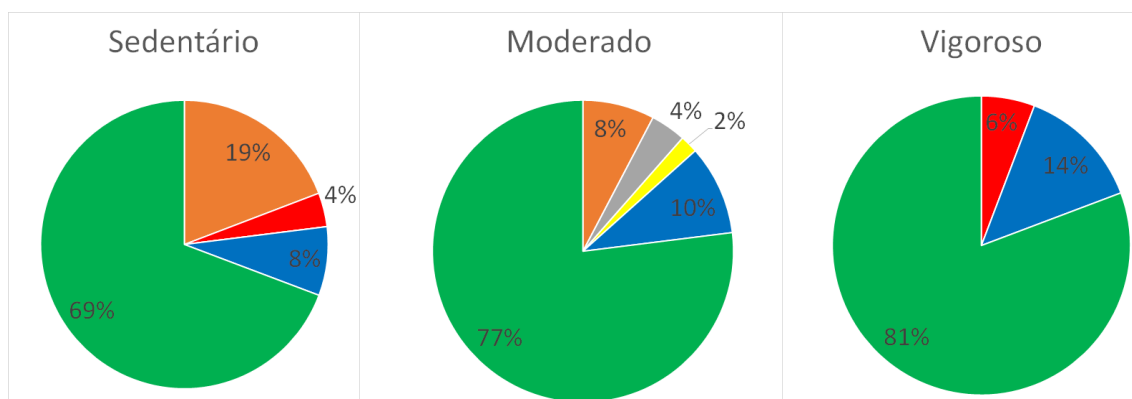


Figura 24 - Gráfico de confusão por tipo de movimento para o Observador 3 sem parâmetros

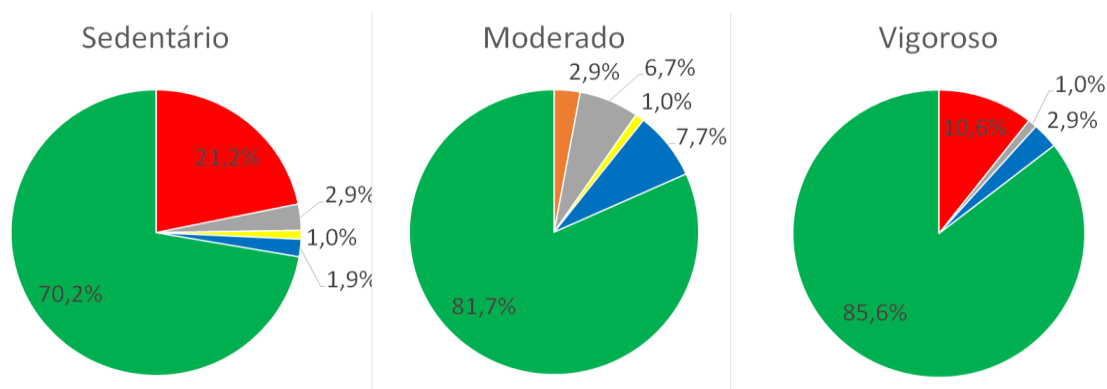


Figura 25 - Gráfico de confusão por tipo de movimento para o Observador 3 com parâmetros

Como conclusão, verifica-se que todos observadores tem uma taxa de confusão alta, o que leva a uma caracterização incorreta em cada tipo de movimento. No caso do movimento vigoroso, como seria de esperar a maior confusão é com o movimento moderado e na contagem do número de pessoas observadas, por ser um tipo de movimento onde os movimentos são rápidos. No caso do movimento sedentário existe maior confusão com o movimento moderado porque a diferença entre os dois é pequena, como por exemplo a passagem de parado para andar. De forma geral estes tipos de erros eram de esperar, pois deve-se a subjetividade inerente a este tipo de caracterização, visto que cada observador tem maneiras diferentes de diferenciar os 3 tipos de movimentos, devidas as características cada movimento serem muito próximos o que dificulta a diferenciação.

4.3.2. CLASSIFICAÇÃO MOVIMENTO A PARTIR DE UM PAINEL DE OBSERVADORES

Uma vez que a observação direta é realizada por humanos, existe um grau de subjetividade inerente. No processo de anotação foi pedido aos observadores para indicarem o seu nível de confiança para cada observação, para assim ser possível perceber a percepção do observador em relação à análise que realizou. A confiança de cada observação nas diferentes fases é um campo que ajuda ao entendimento do nível de certeza existente na caracterização feita.

A confiança teve um fator determinante quando é comparado as observações de cada observador, entrando como um fator de penalização para cada conjunto de observações

Foi determinado duas formas de seleção da melhor observação:

- No caso de dois ou mais observadores concordam, a observação seria igual à dos observadores que concordam;
- Nenhum concorda: Assume-se como correto a observação do observador com um nível de confiança mais alto.

A confiança global em cada observação foi determinada seguindo a equação:

$$C_f = \frac{\sum_i^{N_c} C_i^c + \sum_j^{N_d} (1 - C_j^d)}{N} \quad (1)$$

$$N = N_c + N_d \quad (2)$$

Onde C_i^c é a soma da confiança dos observadores que concordam, C_j^d é soma da confiança dos que discordam para uma a respetiva penalização; N é o número total de observadores.

Tabela 7 – Confiança por movimento e tipo de observação.

Tipo de Observação	Tipo de Movimento	Dois Observadores Concordam (%)	Todos Concordam (%)	Nenhum Concorda (%)
Sem paragens	Sedentário	60	60	-
	Moderado	60	60	-
	Vigoroso	62.5	60	-
Com Paragens	Sedentário	52.5	70	55
	Moderado	55	75	60
	Vigoroso	55	65	57.5

De maneira a ser possível verificar quais os movimentos em que os observadores têm mais confiança, foi criada a Tabela 7. Esta é constituída pela confiança global por movimento. Mais uma vez os observadores, demonstram que a confiança é mais alta quando estão os três de acordo. Em relação ao tipo de observação, os observadores têm maior tendência a concordarem entre si no tipo sem paragens, isto porque não existe nenhuma observação. Estes resultados não vão de encontro ao esperado. Pois era esperado na observação com paragens os observadores terem mais confiança e acertarem mais entre si, isto porque estão sujeitos a menos pressão, não tem tempo para realizar a anotação logo teriam mais tempo para pensar e mais confiança.

5. ABORDAGEM E TESTE PARA CARACTERIZAÇÃO AUTOMÁTICA

Após as análises de todos os dados e verificadas todas as limitações da observação direta, foi demonstrado a importância que uma caracterização automática poderá trazer a este campo de estudo, isto porque poderá ter a capacidade de retirar algumas limitações e subjetividades que o método tradicional tem. Para isso foi necessário realizar um estudo e um conjunto de testes para verificar a viabilidade de um classificador automático.

Inicialmente foi necessário simular o output típico de um algoritmo de seguimento, foi construído um programa que faria a criação de informação de referencia orientada à classificação automática, ou seja, um programa que manualmente realizasse anotações de todos os movimentos e outras características a retirar do vídeo para um output. Seguidamente foi realizado diferentes testes para criação de um classificador automático com a biblioteca LIBSVM.

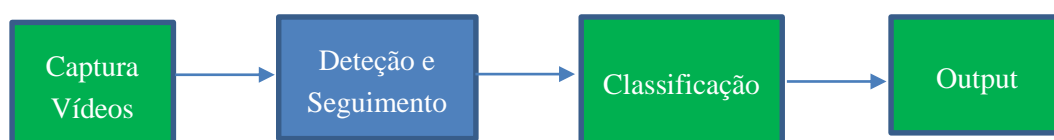


Figura 26 – Processo de criação da classificação automática

Na Figura 26 é demonstrado a sequência lógica para a criação da classificação automática. Nesta dissertação foi realizada a captura de vídeos de pessoas, classificação por observação direta e automática e output.

5.1. SIMULAÇÃO DE UM ALGORITMO DE SEGUIMENTO

No âmbito do estudo do processamento automático de informação foi necessário realizar anotações dos vídeos para a realização da caracterização da atividade física. A anotação

consiste na importação das anotações da observação de referência de modo a simular um algoritmo de seguimento, foi necessário a criação de uma aplicação que fosse capaz da realização dessa tarefa.

Foi concluído que a aplicação teria de ter as seguintes capacidades:

- localização em cada *frame* da posição de cada pessoa;
- diferenciação da orientação da pessoa (deitada/pé);
- do tipo de movimento;
- identificação de cada pessoa;
- número de *frame*.

A aplicação foi desenvolvida através do ambiente gráfico QT Creator que faz parte do projeto Qt que consiste num *framework* QT que é multiplataforma para o desenvolvimento de interfaces gráficas na linguagem C++.

5.2. DESENVOLVIMENTO DA APLICAÇÃO

A aplicação de anotação da atividade física foi desenvolvida com o pensamento de fácil e rápida interação com o utilizador, isto é com a utilização do teclado e rato. O rato através do clique pode definir a posição do objeto em análise na imagem ou seja, retirar diferentes dados, como o tamanho da pessoa ou a sua posição atual (deitado, em pé), sendo representado recorrendo a um caixa envolvente (*Bounding box*) de cada pessoa. Ao teclado foi atribuído diferentes teclas com diferentes funções, tais como diferentes movimento, orientação da pessoa e possibilidade de percorrer o vídeo por *frames*. Todo o trabalho de anotação é guardado num ficheiro de formato *eXtensible Markup Language* (XML).

Funções e funcionamento da aplicação:

- Dois cliques, a *bounding box* é desenhada com um clique no canto superior esquerdo e outro no canto inferior direito da pessoa a anotar;
- Teclas de “1” a “3” do teclado atribui o tipo de movimento (Sedentário/Moderado/Vigoroso);
- Tecla “0” e “9” atribui a orientação 0° e 90° (deitado ou em pé);
- “ESC” para sair da aplicação e ser guardado todo a anotação num ficheiro XML;
- “SPACE” para guardar a anotação em memória;

- “Q”/”W”/”E” usado para avançar *frames* do vídeo, 1/10/30 *frames*;
- “A”/”S”/”D” recuar *frames* do vídeo, -1/-10/-30 *frames*.
- H para interpolação de *frames*.

Realizada a anotação de cada jogador o resultado é guardado num ficheiro XML que contém todas as diferentes *tags* necessárias para identificação da pessoa e seu movimento.

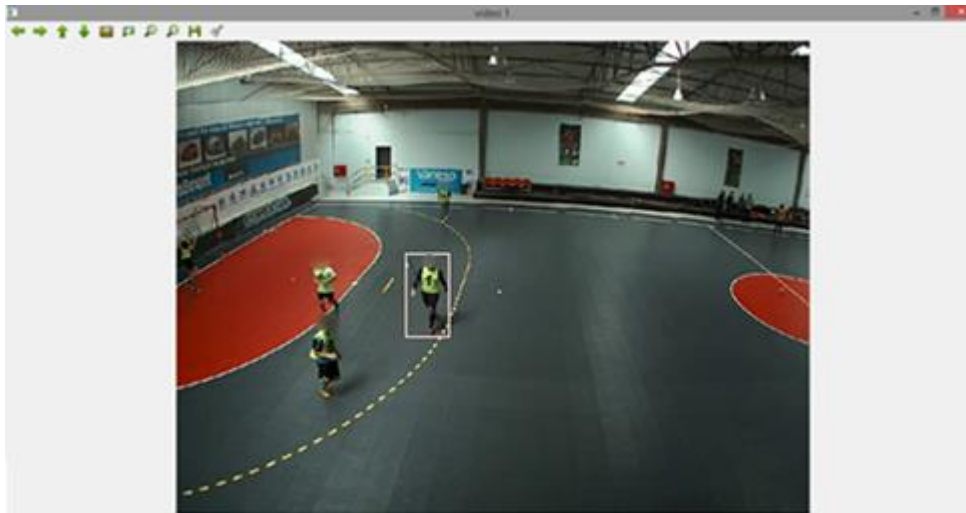


Figura 27 – Exemplo da aplicação.

```

1  <?xml version="1.0" ?>
2  <dataset name="anotacao">
3    <frame number="0">
4      <objectlist/>
5    </frame>
6    <frame number="1">
7      <objectlist>
8        <object id="1">
9          <orientation>90</orientation>
10         <box h="137" w="91" xc="96" yc="324"/>
11         <soplay>1</soplay>
12         <sofit>0</sofit>
13       </object>
14       <object id="2">
15         <orientation>90</orientation>
16         <box h="85" w="73" xc="153" yc="251"/>
17         <soplay>1</soplay>
18         <sofit>0</sofit>
19       </object>
20       <object id="3">
21         <orientation>90</orientation>
22         <box h="140" w="89" xc="44" yc="348"/>
23         <soplay>1</soplay>
24         <sofit>0</sofit>
25       </object>
26       <object id="4">
27         <orientation>90</orientation>
28         <box h="129" w="84" xc="215" yc="270"/>
29         <soplay>1</soplay>
30         <sofit>0</sofit>
31       </object>
32       <object id="5">
33         <orientation>90</orientation>
34         <box h="113" w="62" xc="206" yc="240"/>

```

Figura 28 – Excerto do ficheiro output XML.

Na Figura 28, é possível identificar as diferentes *tags*, cada uma tem como objetivo ajudar no processo de *machine learning*. O ficheiro vai conter: as *frames* existentes no vídeo, a

altura e comprimento de cada *bounding box*, o seu centróide, a orientação da pessoa (0° ou 90°), é atribuído a cada pessoa uma identificação (“id”) e é caracterizado o seu movimento com base na informação de referência orientada para a classificação automática, realizada pelos observadores. O tipo de movimento foi diferenciado pela atribuição de números tais como:

1. Sedentário
2. Moderado;
3. Vigoroso;

Foram realizadas 10 anotações de pessoas com diferentes movimentos, em relação as câmaras de modo a ser apresentado um output mais diversificado e poder ser possível combater todas as limitações da caracterização em vídeo.

Todas estas anotações vão contribuir para a realização de um classificador que será usado para a classificação automática da atividade.

5.3. DEFINIÇÃO DO CLASSIFICADOR AUTOMÁTICO

Como foi referenciado anteriormente na criação de um classificador é necessário estabelecer um conjunto de características que possam definir cada classe de um classificador. Para isso, foram usadas as anotações realizadas anteriormente com as seguintes características:

- Largura e altura da *bounding box*;
- Centróide da *bounding box*;
- Orientação da pessoa;
- Tipo de movimento.

Para a seleção das *features* foi necessário definir que tipo de características, no movimento humano e que estão presentes no output típico de algoritmos de *tracking* que podem definir e diferenciar os diferentes movimentos, como:

- Velocidade instantânea entre dois pontos;
- Velocidade média em um determinado intervalo de tempo;
- Orientação;
- Aceleração;
- Etc.

No caso de caracterização da atividade física através do uso de vídeos, as *features* possíveis que podem ser usadas para a classificação com base nos dados das anotações são:

- Velocidade instantânea mínima e máxima em x e y;
- Velocidade média num determinado número de *frames*;
- Aceleração média em x e y num determinado número de *frames*;
- Orientação
- Altura e largura da *Bounding Box*;

5.4. SELEÇÃO DO CONJUNTO DE DADOS DE TREINO E TESTE

No processo de treino e teste foi necessário escolher o tipo de validação para os testes, isto é, foi necessário escolher que dados seriam usados para treino do classificador e dados para teste. O tipo de validação que poderia ajudar na conclusão do melhor tipo de classificador e suas *features* foi o *Cross-Validation* devido ao conjunto de dados ser finito.

A divisão do conjunto de dados foi dividida da seguinte forma, de maneira a seguir o procedimento do *Cross-Validation*. Do conjunto total de 10 anotações correspondente a 10 pessoas a realizar atividade física, foram usados 9 para dados de treino e 1 para teste, num sistema rotacional exemplificado na Tabela 8 que ilustra 10 Experiências com 10 “id” de 10 pessoas diferentes.

Tabela 8 – Seleção dados de teste e treino

	ID	1	2	3	4	5	6	7	8	9	10
Experiências	1	Teste	treino	treino	treino	treino	treino	treino	treino	treino	treino
	2	treino	Teste	treino	treino	treino	treino	treino	treino	treino	treino
	3	treino	treino	teste	treino	treino	treino	treino	treino	treino	treino
	4	treino	treino	treino	teste	treino	treino	treino	treino	treino	treino
	5	treino	treino	treino	treino	teste	treino	treino	treino	treino	treino
	6	treino	treino	treino	treino	treino	teste	treino	treino	treino	treino
	7	treino	treino	treino	treino	treino	treino	teste	treino	treino	treino
	8	treino	treino	treino	treino	treino	treino	treino	teste	treino	Treino
	9	treino	treino	treino	treino	treino	treino	treino	treino	teste	Treino
	10	treino	treino	treino	treino	treino	treino	treino	treino	treino	teste

Com este método é possível realizar o teste com todas as amostras e assim calcular uma média e desvio padrão, de maneira a analisar a seleção das melhores *features* e intervalo de tempo para a criação do classificador com melhor precisão.

5.5. TREINO E TESTE DO SISTEMA DE CLASSIFICAÇÃO

Para treino e teste do classificador foi necessário estabelecer um conjunto de *features* a usar para a criação do classificador e analisar o seu output de maneira a ser possível encontrar o melhor classificador. Foram usados 3 conjuntos de *features* diferentes em 3 intervalos de tempo.

Intervalo de tempo: 300/150/90 *frames*;

Features:

- Primeiro classificador inclui todas *features* possíveis:
 - Velocidade mínima e máxima em X e Y do intervalo de tempo;
 - Velocidade e aceleração média do intervalo de tempo;
 - Orientação;
 - Altura e largura da *Bounding Box*.
- Segundo classificador:
 - Velocidade mínima e máxima em X e Y do intervalo de tempo;
 - Velocidade e aceleração média do intervalo de tempo;
 - Orientação.
- Terceiro classificador:
 - Velocidade mínima e máxima em X e Y do intervalo de tempo;
 - Velocidade média do intervalo de tempo;
 - Orientação;
 - Altura e largura da *Bounding Box*.

Tabela 9 – Resultados da classificação com *kernel* Linear

Número <i>frames</i>	Tipo de <i>features</i>	Media Precisão (%)	Desvio Padrão (%)	Diferença Máximo e Mínimo (%)
300	1º Classificador	41,57	12	41
	2º Classificador	54,30	14	45
	3º Classificador	42,04	12	41
150	1º Classificador	33,75	12	42
	2º Classificador	60,48	9	28
	3º Classificador	36,65	23	83
90	1º Classificador	30,44	7	19
	2º Classificador	56,04	8	21
	3º Classificador	27,64	7	20

A Tabela 9 apresenta os resultados obtidos com um *kernel* linear e pode-se analisar a média de acerto com os diferentes tipos de *features* e o seu desvio padrão. Analisando este tipo de classificador verifica-se que a taxa de acerto é muito baixa ou seja, a estratégia aplicada, *features* e o *kernel* linear, não é o, mais correto, para este tipo de dados e classificação.

Seguidamente foi realizado o teste com um *kernel* não linear, (Tabela 10), fazendo uma análise e comparação com os resultados obtidos com *kernel* linear e um *kernel* não linear. E nesta análise foi possível concluir que o tipo de classificador para esta problemática terá de usar um *kernel* não linear isto porque as taxas de acerto entre os dois tipos de *kernel* são substancialmente diferentes, ou seja o *kernel* não linear aumenta a taxa de acerto em cerca de 30% e em alguns casos 50%.

Tabela 10 - Resultados da classificação com *kernel não Linear*

Número <i>frames</i>	Tipo de <i>features</i>	Média Precisão (%)	Desvio Padrão (%)	Diferença Máximo e Mínimo (%)
300	1º Classificador	72,56	10	36
	2º Classificador	73,10	11	37
	3º Classificador	73,03	11	36
150	1º Classificador	72,76	8	26
	2º Classificador	73,85	8	23
	3º Classificador	73,37	6	20
90	1º Classificador	70,98	7	22
	2º Classificador	71,31	7	18
	3º Classificador	69,94	8	23

Um terceiro teste foi realizado com o objetivo de otimizar o *kernel não linear*, porque com o uso de vídeos existe características, (altura e largura da *bounding box*), que poderiam condicionar os resultados obtidos e no qual nos testes anteriores não se tinha tido em conta. Foi analisado todas as caraterísticas usadas e foi verificado que o tamanho da *bounding box* era um elemento que não estaria a ser bem aplicado o que poderia contribuir para o valor do desvio padrão. Foi então realizada uma normalização da altura e largura da *bounding box* e aplicado na velocidade, isto porque na visualização dos vídeos foi verificado que a altura estava a limitar o valor da velocidade nas situações em que a pessoa teria diferentes distâncias com câmara. Um exemplo desta situação, pode-se verificar quando uma pessoa está a realizar um movimento vigoroso mais distante da câmara poderia estar a ter a mesma velocidade de quando está mais perto, mas por se encontrar mais longe, sem a normalização, poderá estar a ter uma velocidade muito inferior devido a ter tamanhos mais pequenos de pixéis.

Para a criação dessa normalização foi usada a seguinte expressão para o cálculo da velocidade:

$$V = \frac{\frac{(d1 - d2)}{h}}{T}$$

Legenda:

- V – Velocidade em um determinado instante;
- d1 e d2 – Coordenadas do centróide da *bounding box* em dois determinados pontos;
- T – Tempo entre os dois pontos;
- H – media da altura da *bounding box* nos dois pontos.

Com a normalização da altura da *bounding box* foi realizado um novo treino e teste com o *kernel* anterior.

Tabela 11 - Resultados da classificação com *kernel não* Linear com normalização

Número <i>frames</i>	Tipo de <i>features</i>	Média Precisão (%)	Desvio Padrão (%)	Diferença Máximo e Mínimo (%)
300	1º Classificador	70,62	12	47
	2º Classificador	65,13	10	36
	3º Classificador	63,20	9	27
150	1º Classificador	71,49	9	29
	2º Classificador	74,37	9	28
	3º Classificador	67,36	8	22
90	1º Classificador	73,48	9	25
	2º Classificador	78,01	6	17
	3º Classificador	72,87	9	27

Na Tabela 11 podemos verificar que em determinadas situações a média da taxa de acerto baixou, mas nas situações como no caso do 90 *frames* com aceleração subiu substancialmente e o desvio padrão no geral é mais baixo, isto é, em cada classificação a taxa de acerto é mais constante.

5.6. ANÁLISE DOS RESULTADOS

Na análise de todos os resultados dos testes realizados verifica-se que o *kernel não* linear com normalização da *bounding box* é o que apresenta melhores resultados. A taxa de acerto é alta em cada classificação, em comparação com a observação de referência.

Tabela 12 – Cruzamento dos melhores resultados de todos classificadores

Número <i>Frames</i>	Tipo de <i>features</i>	Tipo de <i>kernel</i>	Média Precisão (%)	Desvio Padrão (%)	Diferença Máximo e Mínimo (%)
300	1º Classificador	Ñ Linear s/ normalização	72,56	10	36
	2º Classificador	Ñ Linear s/ normalização	73,10	11	37
	3º Classificador	Ñ Linear c/ normalização	73,03	11	36
150	1º Classificador	Ñ Linear s/ normalização	72,76	8	26
	2º Classificador	Ñ Linear s/ normalização	73,85	8	26
	3º Classificador	Ñ Linear s/ normalização	73,03	6	20
90	1º Classificador	Linear c/ normalização	73,48	9	25
	2º Classificador	Linear c/ normalização	78,01	6	17
	3º Classificador	Linear c/ normalização	72,87	9	27

Na Tabela 12 é realizada a comparação dos melhores resultados entre todos os testes e é verificado que o melhor classificador é o de 90 *frames* de intervalo de tempo, com kernel não linear e com a normalização da bounding box. Em relação às *features* que possibilitam melhor classificação é verificado que a altura e largura da *bounding box* são um fator limitador na classificação automática. Numa comparação de classificadores com o mesmo intervalo de tempo é possível constatar que nos diferentes tipos de *kernel*, os seus valores foram melhorando.

Em relação ao melhor classificador, foram analisadas as anotações e foi verificado que com 90 *frames* (3 segundos) existe uma menor variação do tipo de movimento da pessoa, isto é, no método tradicional de observação direta é realizado um varrimento do campo de visão da esquerda para a direita, o que não é possível na anotação através de vídeos, porque uma pessoa no início dos 10 segundos está parada, (mais a esquerda do ecrã). Esta situação não implica que a pessoa não comece a movimentar-se, pois sendo assim a *bounding box* vai ter valores diferentes, neste sentido a *bounding box* vai acompanhar a pessoa, logo não se pode desprezar que a pessoa está a realizar um movimento, o que não acontece na observação tradicional, que despreza a pessoa a partir do momento que é

realizado o varrimento, o que leva a uma perda de dados. Sendo esse um dos possíveis motivos para o classificador com o intervalo de tempo de 90 *frames*, ser o que tem melhores resultados. Estes devem-se ao intervalo de tempo ser pequeno o que leva a uma menor variedade de movimentos, bem como uma menor variação do centróide e dos tamanhos da *bounding box*.

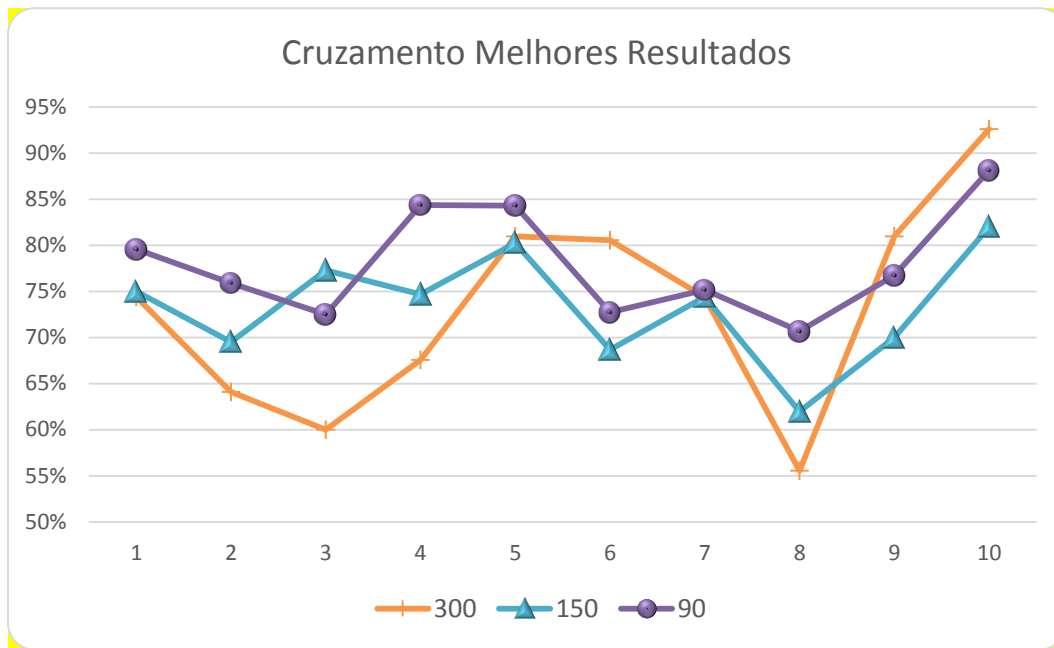


Figura 29 – Gráfico cruzamento melhores classificadores

Na Figura 29 é representado o conjunto total de resultados dos melhores classificadores para cada intervalo de tempo, estes são o conjunto de resultados de teste da classificação de 10 pessoas, onde é possível verificar que o classificador com menor variação de resultados, com mais estabilidade é o de intervalo de tempo de 90 *frames*. Apesar de não ser o classificador que conseguiu o melhor resultado individual, em média é o classificador que menos variação tem, isto é a diferença entre o valor mínimo e máximo é o mais pequeno.

5.7. COMPARAÇÃO ENTRE A CLASSIFICAÇÃO AUTOMÁTICA E A CLASSIFICAÇÃO TRADICIONAL

Após a escolha do melhor classificador possível foi feita uma última consolidação de resultados, através de uma comparação dos resultados da classificação automática e a classificação tradicional. Neste sentido concretizou-se um mapeamento do melhor classificador, de cada intervalo de tempo de 5 das 10 pessoas, ou seja, e comparou-se um

conjunto de pessoas e não por pessoas individuais (SOPLAY). O objetivo do mapeamento é verificar a taxa de acerto dos classificadores aquando comparados com a informação de referência orientada à observação direta e por classificação automática.

Informação de referência orientada à observação direta consiste no modelo que foi realizado pelos observadores em conjunto e em consenso. Neste foi visualizado as sequências e os observadores discutiram a classificação mais correta para cada anotação, sendo que podiam voltar atrás no vídeo e visualizá-lo novamente. Os observadores utilizaram o método tradicional de observação direta, ou seja foi visualizado com intervalos de 10 segundos e com paragens, com possibilidade de voltar a visualizar os intervalos caso existisse confusão, os intervalos eram visualizados através do varrimento do ecrã da esquerda para a direita durante 10 segundos.

A informação de referência orientada à classificação automática foi realizada com ajuda dos observadores, com intervalos de tempo inferiores a 10 segundos, isto porque para o seguimento é impossível realizar o varrimento da esquerda para a direita como na observação direta. Na observação direta cada observador faz o varrimento e ignora qualquer mudança movimento que possa existir posteriormente,. Na observação automática, no caso de uma pessoa mais a esquerda do ecrã, (início do varrimento), tenha um movimento sedentário no início do varrimento e posteriormente alterar o seu estado, não vai ignorar essa mudança porque, a *bounding box* vai acompanhar o observado sempre ao contrario da observação direta, logo na observação automática trás uma grande vantagem que é não há perda de dados em qualquer instante. Pode ser vista como mais exigente porque uma pessoa é analisada em todo o intervalo e os intervalos de tempo são mais corretos isto porque na observação direta o vídeo é parado a cada 10 segundos mas manualmente, ou seja vão existir diferenças entres paragens enquanto na observação automática é sempre parado no mesmo número de *frames*.

Seguidamente é apresentado uma comparação entre a classificação automática e por observação direta com os dois modelos de referência.

Tabela 13 – Comparação do acerto entre Observação Tradicional e Automática com os dois modelos de referência.

		Comparação modelo referencia	Nº Pessoas com movimento erradamente classificado					
			0	1	2	3	4	5
Taxa de Previsão	Orientada à Observação Direta	Classificação Automática	6,67%	20,00%	43,33%	13,33%	6,67%	10,00%
		Classificação Observadores	26,67%	33,33%	26,67%	6,67%	6,67%	0,00%
	Orientada à Classificação Automática	Classificação Automática	26,67%	50,00%	23,33%	0,00%	0,00%	0,00%
		Classificação Observadores	6,67%	26,67%	23,33%	20,00%	13,33%	10,00%

Na Tabela 13 está ilustrada a percentagem de acerto de cada classificação, isto é, a quantas pessoas o classificador acertou em cada anotação, por exemplo a classificação automática em comparação com a informação de referência orientada à observação direta acertou em 6.67% das anotações em 5 pessoas e 20% em 4, ou seja, falhou a classificação de uma pessoa em 20% das anotações. O classificador automático usado foi o que contém o intervalo de tempo mais próximo da observação direta ou seja os 300 *frames*/segundo (10 segundos).

Cada anotação da classificação dos observadores foi realizada pela concordância entre eles, ou seja:

- Se dois ou mais observadores concordam a anotação é igual à sua concordância;
- Nenhuma concordância entre os três - é escolhida a anotação do observador com maior confiança.

Em relação aos resultados da comparação com a informação de referência orientada à observação direta, verifica-se que apesar do classificador automático ter uma percentagem de acerto nas 5 pessoas baixa, tem uma maior percentagem de acerto, 76,67%, onde acerta pelo menos 2 pessoas contra 66,67% da classificação dos observadores. Neste caso apesar da observação direta ter um maior acerto na classificação das 5 pessoas, também em mais de metade dos casos falha pelo menos uma pessoa. O maior acerto nas 5 pessoas da observação direta explica-se pela grande influência e perda de dados que existe na classificação por observadores, em relação ao classificador automatico.

Em relação a comparação das classificações mas com a informação de referência orientada ao classificador automático o que se destaca é o classificador automático que falha no máximo 2 pessoas.

Comparando os dois classificadores com as duas informações de referência, podemos verificar que o classificador automático tem melhorias substanciais com o modelo de referência automático como era de esperar, mas de um modo geral o classificador automático tem melhores resultados médios no total dos dois classificadores.

Tabela 14 – Média do acerto dos classificadores com os dois modelos de referência

		Nº Pessoas com movimento erradamente classificado					
		0	1	2	3	4	5
Somo total 2 classificadores	Classificação Automática	16,67%	35,00%	33,33%	6,67%	3,33%	5,00%
	Classificação Observadores	16,67%	30,00%	25,00%	13,33%	10,00%	5,00%

Na Tabela 14 foi somado o acerto dos dois classificadores nos dois modelos de referência, na mesma é possível verificar que o classificador automático tem maior acerto em mais de metade das pessoas, (3, 4 ou 5 pessoas), ou seja, em 85% das anotações acerta pelo menos em 3 pessoas por anotação, contrariamente ao obtido da classificação por observação direta (71,67%).

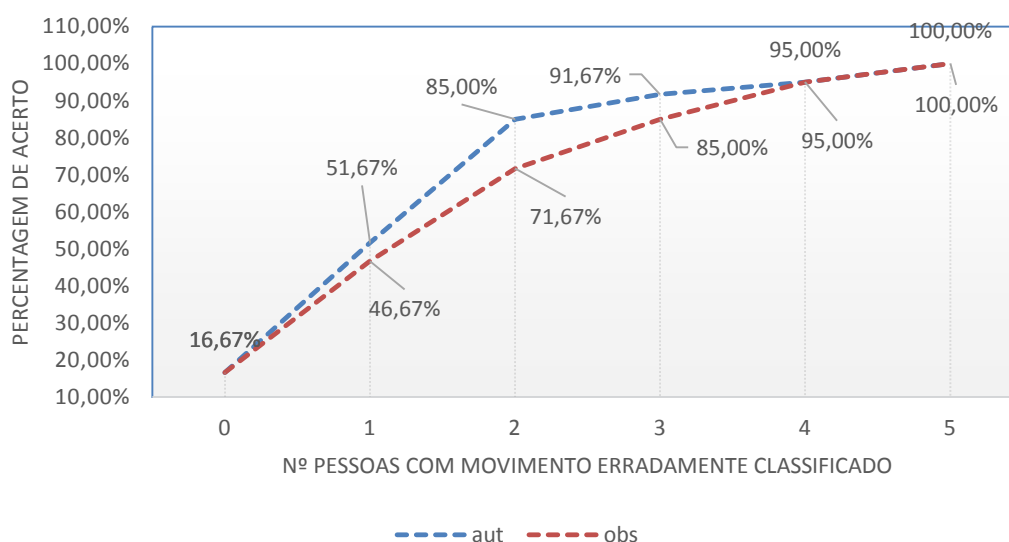


Figura 30 – Gráfico de tendência acumulada da média dos dois classificadores nos dois modelos de referência.

Na Figura 30 podemos verificar que o classificador automático atinge o seu pico mais acentuado de acerto, quando falha no máximo 2 pessoas, (85%), e tende a estabilizar, enquanto o classificador por observação direta está bastante abaixo (71,67%), por conseguinte, constata-se que este tem tendência a falhar mais pessoas nas suas classificações.

6. CONCLUSÕES E DESENVOLVIMENTOS FUTUROS

O trabalho apresentado nesta dissertação foi realizado no contexto dos métodos de caracterização da atividade física, tendo incidido particularmente na análise de processos de melhoria das alternativas SOPLAY e SOFIT utilizando técnicas de processamento automático de imagens de vídeo através de *Machine Learning*.

Na primeira etapa, realizou-se um estudo detalhado das metodologias atualmente utilizadas nos processos de caracterização de atividade física, identificaram-se importantes limitações que os diferentes métodos apresentavam.

Uma das conclusões mais evidentes foi a verificação que diferentes fatores externos tornam a observação direta subjetiva e não linear, ou seja, diferentes observadores apresentam anotações diferentes e, cada um deles tem desempenhos diferentes ao longo do processo completo.

Esta conclusão permitiu validar a ideia inicial de que a utilização de novas metodologias ou tecnologias poderiam conduzir a uma otimização deste processo. A análise do estado da arte e das bibliotecas de *software* existentes em diversas plataformas levou à decisão de utilizar técnicas de *Machine Learning*, nomeadamente SVM, para a implementação de novas abordagens.

A solução desenvolvida implicou a captura de diferentes sequências de vídeo para permitir validar a aproximação utilizada de forma não enviesada e influenciada pelo tipo de conteúdo. Os testes conduzidos após a criação de um modelo de um classificador automático testado com diferentes tipos de parâmetros, (velocidade, aceleração, orientação da pessoa a classificar) e condições, (diferentes intervalos de tempo), permitiu

demonstrar a capacidade de gerar resultados comparáveis ou melhores que a observação direta, com diversos benefícios:

- não enviesamento;
- independente do cansaço dos observadores ao longo das anotações;
- objetividade;
- capacidade de analisar grandes quantidades de informação.

Deste modo, conclui-se que os objetivos inicialmente propostos foram cumpridos e que este trabalho pode contribuir para a resolução do problema existente na atualidade relativamente à execução da atividade física.

A realização de testes mais exaustivos com diferentes sequências de vídeo, que incluam diversos cenários e faixas etárias, deverá ser uma das etapas futuras de forma a garantir a universalidade da abordagem.

No âmbito desta dissertação não foi avaliada de que forma o comportamento dos observados pode ser afetado pela presença de observadores no local, como também a fusão de diferentes sensores (GPS, acelerómetros, etc.). Esta análise pode revelar-se também importante.

A introdução de um algoritmo de seguimento de pessoas no *software* de classificação automática deverá ser avaliada com vista à redução do impacto do ruído na classificação. É também previsível que isto permita reduzir o tempo necessário tornando assim todo o processo mais eficiente. Por fim, a análise de outras características para a criação do classificador pode permitir também melhorias.

Referências Documentais

- [1] Fulton JE, McGuire MT, Caspersen CJ, Dietz WH. Interventions for weight loss and weight gain prevention among youth: current issues. *Sports Med.* 2001.
- [2] Das P, Horton R. Rethinking our approach to physical activity. *Lancet.* Jul 21 2012.
- [3] Sallis J, Owen N. *Physical Activity & Behavioral Medicine.* Vol 3: Sage Publications; 1999.
- [4] Sallis JF, Kraft K, Linton LS. How the environment shapes physical activity: a transdisciplinary research agenda. *Am J Prev Med.* Abril de 2002.
- [5] Dishman RK, Motl RW, Saunders R, et al. Self-efficacy partially mediates the effect of a school-based physical-activity intervention among adolescent girls. *Prev Med.* Maio de 2004.
- [6] Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW. Correlates of physical activity: why are some people physically active and others not? *Lancet.* Julho de 2012.
- [7] Página de Internet – 13 de Julho de 2013, <http://dapa-toolkit.mrc.ac.uk/physical-activity-assessment/methods/observation/index.html>
- [8] McKenzie TL, Crespo NC, Baquero B, Elder JP. Leisure-time physical activity in elementary schools: analysis of contextual conditions. *J Sch Health.* Outubro de 2010.
- [9] Kohl HW, 3rd, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. *Lancet.* Julho de 2012.
- [10] Pratt M, Sarmiento OL, Montes F, et al. The implications of megatrends in information and communication technology and transportation for changes in global physical activity. *Lancet.* Julho de 2012.
- [11] McKenzie TL. 2009 C. H. McCloy Lecture. Seeing is believing: observing physical activity and its contexts. *Res Q Exerc Sport.* Junho de 2010.
- [12] Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet.* Julho de 2012.
- [13] Sallis JF, Taylor WC, Dowda M, Freedson P, Pate R. Correlates of Vigorous Physical Activity for Children in Grades 1 Through 12: Comparing Parent-Reported and Objectively Measured Physical Activity. *Ped.Exerc.Science.* 2002.
- [14] Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc.* Agosto de 2011.

- [15] Heath GW, Parra DC, Sarmiento OL, et al. Evidence-based intervention in physical activity: lessons from around the world. *Lancet*. Julho de 2012.
- [16] McKenzie TL. SOFIT - System for Observing Fitness Instruction Time. 2002.
- [17] McKenzie TL. SOPLAY - System for Observing Play and Leisure Activity in Youth. 2002.
- [18] Bar-Or, O., Rowland T. *Pediatric Exercise Medicine: From physiologic principles to health care application*. Human Kinetics. Champaign, IL. 2004
- [19] Schmidt, R., Wrisberg, C. *Motor Learning and Performance*. Human Kinetics. Champaign, IL. 2000
- [20] Página de Internet – 15 de Abril de 2014, <http://dapa-toolkit.mrc.ac.uk/physical-activity-assessment/methods/observation/index.html>.
- [21] Ralph Maddison, Cliona Ni Mhurchu. Global positioning system: a new opportunity in physical activity Measurement. Novembro de 2009.
- [22] Mitch J. Duncan, Hannah M. Badland, W. Kerry Mummery. Applying GPS to enhance understanding of transport-related physical activity. Outubro de 2008.
- [23] Richard Troiano, PhD. Physical Activity Assessment Using Accelerometers.
- [24] Página de Internet – 15 de Abril de 2014, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3292184/>
- [25] Página de Internet – 13 de Maio de 2014, [2013http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/](http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/)
- [26] Alper Yilmaz, Omar Javed, Mubarak Shah. Object Tracking: A Survey. 2006.
- [27] Luís F. Teixeira, Jaime S. Cardoso, and Luís Corte-Real. Object segmentation using background modelling and cascaded change detection. *Journal of Multimedia (JMM)*, 2007.
- [28] Cristopher Wren, Ali Azarbayejani, Trevor Darrell, and Alex Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 1997.
- [29] Zhenjun Han, Qixiang Ye, and Jianbin Jiao. Combined feature evaluation for adaptive visual object tracking. *Computer Vision and Image Understanding*, January 2011.
- [30] Tao Zhao. Model-Based Segmentation and Tracking of Multiple Humans in Complex Situations. PhD thesis, Faculty of the Graduate School of the University of Southern California, 2004.
- [31] Ismail Haritaoglu, David Harwood, and Larry S. Davis. W4: Real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, August 2000
- [32] Nils T. Siebel and Steve J. Maybank. Fusion of multiple tracking algorithm for robust people tracking. In *Proceedings of European Conference on Computer Vision*, 2002.

- [33] Darius M. Gavrilă. The visual analysis of human movement: A survey. Computer Vision and Image Understanding: CVIU, 1999.
- [34] Abby C. King, Dan Stokols, Emily Talen, Glenn S. Brassington, Richard Killingsworth. Theoretical Approaches to the Promotion of Physical Activity. 2002.
- [35] Rob Schapire, Theoretical Machine Learning, 2008.
- [36] Página de Internet – 13 de Maio de 2014, [https://www.projectrhea.org/oldkiwi/index.php/Lecture_21_-_Decision_Trees\(Continued\)](https://www.projectrhea.org/oldkiwi/index.php/Lecture_21_-_Decision_Trees(Continued))
- [37] Página de Internet – 13 de Maio de 2014, <http://www.mlplatform.nl/what-is-machine-learning/>.
- [38] Christopher J.C. Burges. A tutorial on Support Vector Machines for Pattern Recognition. 1998.
- [39] Página de Internet – 15 Setembro de 2014, <http://crsouza.blogspot.pt/2010/03/kernel-functions-for-machine-learning.html>
- [40] Página de Internet – 15 Setembro de 2014, <http://www.quora.com/What-are-Kernels-in-Machine-Learning-and-SVM>
- [41] Página de Internet – 15 Setembro de 2014, <http://www.support-vector-machines.org/>
- [42] Página de Internet – 15 Setembro de 2014, <https://www.npmjs.org/package/node-svm>
- [43] Página de Internet – 13 de Maio de 2014, <http://www.mblondel.org/journal/2010/09/19/support-vector-machines-in-python/>
- [44] Página de Internet – 13 de Maio de 2014, <http://www.albany.edu/cancer/genomics/faculty/ikuznetsov/research.html>
- [45] Página de Internet – 13 de Maio de 2014, <http://elefant.developer.nicta.com.au/>
- [46] Página de Internet – 13 de Maio de 2014, <http://montepython.sourceforge.net/>
- [47] Página de Internet – 13 de Maio de 2014, http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/data_mining.htm.
- [48] Página de Internet – 09 Junho de 2014, <http://pt.slideshare.net/narendra.sisodiya/python-presentation-presentation>.
- [49] Página de Internet – 09 Junho de 2014, <https://www.python.org/doc/essays/blurb/>
- [50] BRADSKY, G. R.; PISAREVSKY, V.; BOUGUET, J. Learning OpenCV: Computer Vision with the OpenCV Library. Springer, 2006.
- [51] Página de Internet – 15 Setembro de 2014, <http://opencv.org/about.html>
- [52] Chih-Chung, Chih-Jen Lin: LIBSVM: A Library for Support Vector Machines, 2013